The Organization of Health Care Services for Children and Youth: Synthesis of the Evidence to Help Guide the Integration and Consolidation of Pediatric Health Services

Final Report
to
Health Canada

July 2003

Final Report to Health Canada
Health Canada Project (6785-15-7001/4450003)

The Organization of Health Care Services for Children and Youth: Synthesis of the Evidence to Help Guide the Integration and Consolidation of Pediatric Health Services

July 7, 2003

Partner:
Alberta Research Centre for Child Health Evidence, University of Alberta

Authors:
Kelly Russell, Lisa Hartling, Lisa Tjosvold, Ellen Crumley, Terry Klassen

For further information please contact:

Dr. Terry Klassen
Professor of Pediatrics and Chair
Department of Pediatrics
University of Alberta
2C3.67 Walter C. Mackenzie
Health Sciences Centre
Edmonton, Alberta
T6G 2R7
CANADA
email: terry.klassen@ualberta.ca
phone: 780-407-7083
fax: 780-407-8538

Financial contribution from the Health Policy Research Program Health Canada

The views expressed herein do not necessarily represent the official policy of Health Canada

TABLE OF CONTENTS

English summary v
French summary vii
Executive summary ix
Acknowledgements xi
Chapter 1 Introduction 1
Theoretical explanations of the volume-outcome relationship 1
Practice makes perfect 1
Selective referral 1
Regionalization: definition, advantages and disadvantages 1
Definition 1
Advantages 2
Disadvantages 2
Role of case-mix 2
Previous research 3
Chapter 2 Objectives 4
Chapter 3 Methods 5
Identifying the literature 5
Selection of eligible studies 5
Quality assessment 6
Data extraction 6
Data analysis 7
Peer-review 7
Chapter 4 Results 8
Results of literature searches 8
Study characteristics 8
Appendicitis 9
Cancer 10
Cardiac 12
Neonatal: level of care 14
Neonatal: volume or regionalization 16
NICU 19
PICU 21
Transplant 22
Trauma 23
Miscellaneous 24
Findings 26
Overall 26
Clinical areas and role of risk adjustment 26
Appendicitis 26
Cancer 26
Cardiac 27
Neonatal: level of care 27
Neonatal: volume or regionalization 28

NICU 28
PICU 29
Transplant 29
Trauma 30
Miscellaneous 30
Quality of studies and limitations 31
Overall quality 31
Observational studies 31
Limitations with the literature 32
Outcomes measured 32
Objectives studied 33
Definition of volume 33
Hospital versus patient volume 33
Power 34
Chapter 6 Conclusions and recommendations 35
Implications for delivery of health services 35
Implications for research 35
References ..... **Error! Bookmark not defined.**
Appendix I: Tables 46
Appendicitis 46
Cancer 48
Cardiac 55
Neonatal: Level of Care 63
Neonatal: Volume or Regionalization 80
NICU 97
PICU 103
Transplant 108
Trauma 110
Miscellaneous 115
Appendix II: Pathways for the Volume/Outcome Relationship 118
Appendix III: Detailed Quality Table 119
Appendix IV: Literature Search Strategies 123
Appendix V: Inclusion Form 127
Appendix VI: Figure of Inclusion/Exclusion Flow Chart 129
Appendix VII: Down's and Black Quality Assessment 130
Appendix VIII: Data Extraction Form 138

English summary

The volume-outcome relationship was put forth as a possible explanation for variation in health outcomes. This relationship refers to the possibility that the more often a procedure is performed or a type of patient is treated, proficiency and expertise will be gained. In order to accurately determine how volume influences outcome, the patients' risk for poor outcomes must be standardized or adjusted. We systematically reviewed studies that examined how volume or regionalization of health services affected patient outcomes or utilization of health services.

Five electronic databases were searched for potentially relevant studies. Studies were also identified from authors and reference lists of included studies and conference proceedings. Two reviewers independently screened the search output and applied the following inclusion criteria: 1) study evaluated a volume-outcome relationship or impact of service regionalization; 2) study had a comparison group; 3) study measured objective outcomes (e.g., mortality); and 4) participants were providers of pediatric health services. Using Downs and Black's checklist and the degree of risk adjustment for patient's risk of poor outcome, quality was assessed independently by two reviewers. Data was extracted by one reviewer and checked for accuracy and completeness by a second. Due to substantial heterogeneity, meta-analysis was not possible.

One hundred-six studies were included and covered the following clinical areas: appendicitis (n=2), cancer (11), cardiac (12), level of care (23), neonatal (28), NICU (8), PICU (7), transplant (3), trauma (7), and miscellaneous (5). The median quality index was 21/29 (inter-quartile range 18,22). There was inconclusive evidence for a volume-outcome relationship for appendicitis management. Survival in children diagnosed with cancer tended to increase when treated at high volume, pediatric oncology treatment centres; this was dependent on cancer type. Volume was also positively associated with outcome for children requiring organ transplants. A volume-outcome relationship existed for children undergoing a breadth of cardiac surgeries; outcomes also improved with regionalization. Increased level of care resulted in improved survival for low birth weight and high-risk neonates. In NICU and PICU settings, volume was not related to improved outcome, however patients treated in institutions with a high level of resources had better health outcomes. Increased volume did not improve outcome in the trauma settings; it is unclear if increased resources or treatment in pediatric trauma centres affects patient outcome.

The volume-outcome relationship appears to exist in most settings examined, but may be dependent on the procedure or the particular pediatric specialty. Children with a high risk of dying appear to fare better in high volume and/or high resource level of hospital. Critically and severely ill pediatric population and premature neonates have improved outcomes when treated in resource-intensive hospitals. The outcome for surgical procedures is generally more favorable in high volume hospitals. However, this does not apply to all surgical procedures and is more evident in complex and rare operations.

104-03569/

Because resource availability appears to be a larger predictor of health outcome than volume, severely ill children should be treated in institutions with an increased level of resources. Finally, the outcome for high-risk pregnancies and premature babies is improved when delivered in Level III hospitals. For trauma management and delivery of babies, a well-developed transportation system is required to shuttle patients to the hospital that matches their required level of resource. There is currently insufficient evidence to quantify a volume threshold for care in different centers.

French summary

Afin de tenter d'expliquer les fluctuations des résultats pour la santé, l'idée a été émise qu'il y avait un rapport entre le volume et le résultat. Ce rapport fait référence au fait que plus souvent on pratique une intervention ou soigne un type de patient, plus on gagnera en compétence et en expertise. Afin de déterminer précisément comment le volume influence le résultat, le risque des patients d'obtenir de médiocres résultats doit être normalisé ou ajusté. Nous avons systématiquement passé en revue les études qui analysaient la façon dont le volume ou la régionalisation des services de santé affectait les résultats du patient ou l'utilisation de ces mêmes services de santé.

Nous avons effectué des recherches dans cinq bases de données électroniques pour trouver des études qui pouvaient s'avérer pertinentes. Nous avons identifié les auteurs de ces études ainsi que les listes de références qui se trouvaient comprises dans ces études et dans les résultats de conférences. Deux examinateurs ont, chacun de leur côté, analysé de très près les données soumises à la recherche et ont appliqué les critères d'inclusion suivants :1. l'étude évaluait le rapport entre le volume et le résultat ou l'impact de la régionalisation des services ; 2. l'étude présentait un groupe de comparaison ; 3. l'étude quantifiait des résultats objectifs (p.ex. mortalité) et 4. les participants offraient des soins de santé pédiatriques. Deux examinateurs ont séparément, évalué la qualité : ils se sont servis de la liste de contrôle de Downs et Black et du degré d'ajustement au risque en ce qui concerne le risque du patient d'obtenir de médiocres résultats. Le premier examinateur s'est chargé d'extraire les données pendant que le deuxième se chargeait d'en vérifier la précision et l'exhaustivité. Une méta-analyse n'était pas envisageable à cause de la trop grande hétérogénéité des données.

Nous avons inclus cent six études qui couvraient les services cliniques suivants : l'appendicite (2), le cancer (11), les problèmes cardiaques (12), le niveau des soins (23), la néonatalogie (28), l'unité néonatale des soins intensifs (8), l'unité pédiatrique de soins intensifs (7), les greffes (3), la traumatologie (7) et divers autres services (5). L'indice de qualité moyen se situait à 21/29 (l'intervalle interquartile étant de 18,22). En ce qui concerne la gestion de l'appendicite, aucune preuve concluante n'a démontré le rapport entre le volume et le résultat. Le taux de survie des enfants atteints de cancer avait tendance à augmenter lorsqu'ils étaient soignés dans des centres d'oncologie pédiatrique à gros volume. Mais cela dépendait du type de cancer dont ils étaient atteints. Pour les enfants qui avaient besoin d'une greffe d'organes, le volume était directement lié au résultat. Il y avait un rapport entre le volume et le résultat pour les enfants qui devaient subir une diversité de chirurgies cardiaques; les résultats s'amélioraient également avec la régionalisation. Des niveaux de soins élevés avaient pour résultat d'améliorer la survie des nouveaux-nés à haut risque et dont le poids à la naissance était insuffisant. En ce qui concerne l'unité pédiatrique de soins intensifs et l'unité néonatale des soins intensifs, le volume n'était pas relié à de meilleurs résultats, mais les patients soignés dans des hôpitaux à très grandes ressources avaient de meilleurs résultats de santé. Un plus gros volume n'améliorait pas le résultat en traumatologie : il n'est pas clair si de meilleures

ressources ou si des traitements dans des centres de traumatologie pédiatrique influencent le résultat pour le patient.

Le rapport entre le volume et le résultat semble exister dans la plupart des cas de figure analysés, mais pourrait dépendre des interventions ou de la spécialité spécifique du centre pédiatrique. Les enfants dont le risque de mourir est élevé semblent s'en sortir mieux dans un hôpital à gros volume ou à grandes ressources. Les enfants très malades ou gravement atteints ainsi que les nouveaux-nés prématurés ont de meilleurs résultats lorsqu'ils sont soignés dans des hôpitaux exigeants en termes de ressources. Le résultat lors d'interventions chirurgicales est généralement plus favorable dans des hôpitaux à volume élevé. Cela ne s'applique cependant pas à toutes les interventions chirurgicales et est plus manifeste lors d'interventions complexes et rares. Les enfants très malades devraient être soignés dans des hôpitaux qui possèdent un niveau élevé de ressources parce que leur disponibilité semble être un plus grand indicateur de résultat pour la santé que le volume. Enfin, les résultats pour les grossesses à haut risque et pour les prématurés sont améliorés lorsque les naissances ont lieu dans des hôpitaux de niveau III. En ce qui concerne la gestion des traumas et celle de la naissance des bébés, il est nécessaire d'avoir un système de transport bien développé afin d'acheminer les patients vers l'hôpital correspondant au niveau des ressources nécessaires. Il n'y a pour l'instant aucune preuve suffisante pour quantifier le seuil de volume des soins dans différents centres.

Executive summary

Over thirty years ago, the volume-outcome relationship was put forth as an explanation for varying health outcomes. The volume-outcome relationship refers to the concept that the more often a procedure is performed or a type of patient is treated, the physician and hospital will gain proficiency and expertise in their treatment. In an attempt to explain the volume-outcome relationship, two theories have been developed. The "practice makes perfect" theory states that repetition of procedure and type of patient improves outcome. The "selective referral" theory states that patients are referred to physicians with superior outcomes and the increased patient referral increases a physician's volume of patients. In order to accurately determine how volume influences outcome, the patients' risk for poor outcomes must be standardized or adjusted. We systematically reviewed 106 studies that examined how volume or regionalization of health services affected patient outcomes or utilization of health services.

A medical librarian conducted electronic searches in the five following bibliographic databases: Medline, EMBASE, Cinhal, HealthStar, and Web of Science. Two reviewers independently screened the search output and the full text of potentially relevant studies was obtained. Two reviewers applied the inclusion criteria to 1,213 studies. Studies were included if they examined a volume-outcome relationship or evaluated the impact of regionalization of services in a pediatric population. The study must have included a comparison group and have measured objective outcomes such as mortality, length of stay, or admission rates. Third, the participants had to be either: facilities or healthcare workers delivering pediatric health services, or any provider of child health (including governmental, non-governmental, or private organizations). Eighty-seven studies were deemed relevant. An additional ten studies were identified from conference proceedings or reference lists of included studies. Writing to authors of included studies identified another nine studies. Quality was assessed independently by two reviewers. Two techniques were used to evaluate studies for their methodological rigor, Downs and Black's checklist for randomized and non-randomized studies of health care interventions and the degree of risk adjustment for different levels of case-mix. Data was extracted by one reviewer and checked for accuracy and completeness by a second. Due to great amounts of heterogeneity, a quantitative analysis could not be undertaken.

The clinical areas studied were: appendicitis (n=2), cancer (11), cardiac (12), level of care (23), neonatal (28), NICU (8), PICU (7), transplant (3), trauma (7), and miscellaneous (5). The median quality index was 21 (inter-quartile range 18,22), and the majority of studies were conducted in the United States. The quality of the studies improved over time.

There was inconclusive evidence for a volume-outcome relationship for the management of appendicitis. Survival in children diagnosed with cancer tended to increase when they were treated at high volume oncology treatment centres, however this was dependent on the type of cancer being studied. Children also tended to have better outcomes when treated in pediatric oncology centres versus other types of treatment institutions. Volume

was also positively associated with outcome for children requiring organ transplants. A volume-outcome relationship existed for children undergoing a breadth of cardiac surgeries; regionalization of cardiac care also improved health outcomes. Increased level of care resulted in improved survival for low birth weight and high-risk neonates. For NICU and PICU settings, volume was not related to improved outcome, however patients treated in institutions with a high level of resources had better health outcomes than their counterparts treated in basic hospitals. Increased volume did not improve outcome in the trauma settings, however it is unclear if increased resource availability or treatment in pediatric trauma centre affects patient outcome.

The volume-outcome relationship appears to exist in most settings examined, but seems to be dependent on the procedure being performed or the particular pediatric specialty. Children with a high risk of dying appear to fare better in high volume and/or high level of hospital. The critically ill pediatric population (i.e., admission to an intensive care) or severely ill or premature neonates have improved outcome when treated in resource intensive hospitals. There is also a volume-outcome relationship for most pediatric surgical procedures examined.

Based on the evidence reviewed, the outcome for surgical procedures is generally more favorable in high volume hospitals. However, this does not apply to all surgical procedures, and is more evident in complex and rare operations. Because resource availability appears to be a larger predictor of health outcome than volume, severely ill children fare better in institutions with an increased level of resources. Finally, the outcome for high-risk pregnancies and premature babies is improved when delivered in Level III hospitals. For trauma management and delivery of babies, a well-developed transportation system is required to shuttle patients to the hospital that matches their required level of resource.

Acknowledgements

We would like to thank Kris Cramer (Alberta Research Centre for Child Health Evidence -ARCHE) and Jacqueline Page (Cochrane Child Health Field) for their assistance in screening studies for potential relevance. We must thank Natasha Wiebe (ARCHE) for providing statistical guidance and support. Appreciation is extended to Colette Freitas and John Russell (ARCHE) for their administrative support and the collection of articles. Also, we are grateful to the authors of the primary studies who reviewed the preliminary list of included documents in an effort to find additional studies. Thank you to Lisa Simpson for developing the template of the Pathway of the Volume-Outcome Relationship figure. Finally, we are grateful to the experts who provided feedback on the draft report:

- Dr. Anita Kozyrskyj, University of Manitoba, Canada
- Ms. Elaine Orrbine, Canadian Association of Paediatric Health Centres, Canada
- Dr. Lisa Simpson, Agency for Healthcare Research and Quality, USA

Chapter 1 Introduction

The volume-outcome relationship refers to the idea that the more often hospitals or physicians perform a specific procedure or treat patients with a particular condition, the more proficient they will become and thus improve patient outcomes, such as mortality, morbidity, and shorten length of stay. This relationship has been demonstrated in several adult procedures.[1,2]. Since children represent a relatively small proportion of individuals requiring intensive health services, there is often less opportunity for hospitals and health care providers to gain the experience needed to perfect skills specific to pediatric care.[3] If a volume-outcome relationship does exist for pediatric health services, efforts should be made to centralize health services to ensure that physicians and/or hospitals maintain an adequate volume to optimize clinical outcomes. Two theories have been put forward to explain the volume-outcome relationship: "practice makes perfect" and "selective referral".[4]

Theoretical explanations of the volume-outcome relationship

Practice makes perfect

The rationale behind the ubiquitous philosophy of "practice makes perfect" is that the more experience one gains with a specific procedure or treatment, the better they will become. Clinical outcomes may improve for several reasons: the individual clinician may become more skilled at the specific technique, the medical or surgical team may become more cohesive and efficient, or a hospital may invest in more sophisticated equipment because they are conducting a high number of procedures or see many similar patients.[5]

Selective referral

The second explanation of the volume-outcome relationship is the process of selective referral. As a physician or hospital develops its reputation as having exceptional health outcomes, additional patients will be referred to them and subsequently, volume will increase. Patients that are referred to high volume hospitals tend to be less ill and are therefore more likely to have improved health status, regardless of hospital volume.[6] Also, hospitals with superior health outcomes may attract better clinicians.

Regionalization: definition, advantages and disadvantages

Definition

Regionalization refers to an explicit plan to deliver health care to a defined region or catchment area to ensure that patients receive the optimal care they require. Regionalization implies providing a level of care for the patient that matches the patient's level of risk. Regionalization of primary health care involves bringing together inputs, organization, management, and delivery of particular service functions to improve the efficiency and quality of health care delivery.[7] For the purpose of this report, regionalization and centralization are used synonymously.

Advantages

There are several advantages to centralizing health services.[5] Perhaps the most intuitive advantage is the increased opportunity to perform and refine specific procedures and treatments. Greater repetition of procedures may result in a decreased risk of complications and better management plans for dealing with deleterious turn of events. Also, the finest physicians are recruited to institutions that have increased opportunity to perform their skill. In addition, in large hospitals there is increased opportunity for in-house consultations.

There are also economic advantages for regionalizing health services.[5] Regionalization avoids duplication of expensive resources and equipment. The increased volume makes it economically feasible to develop new programs and subspecialties within the hospital. Also, medical research is more likely to be conducted in large hospitals, as high volume allows such centres to recruit more patients into clinical trials. Large volume institutions have an increased opportunity for collecting data and increased capacity for quality assurance.

Disadvantages

Patient care can be negatively affected by regionalization.[5] Patients may be detached from their supportive environment as their care is transferred to a larger centre. An effective and efficient centralized system is also resource intensive. For instance, in order for patients to be treated at their region's main hospital, extensive transportation and support programs for critically ill patients must exist. Without adequate ambulatory transportation to high volume hospitals, centralization will not work. Further, a system to handle extended recovery or long-term care must be in place to transfer patients back to less resource intensive, smaller hospitals.

Regionalized care may improve outcomes in high volume institutions, however it may result in degeneration of skills and mediocrity at lower volume hospitals.[5] If clinicians at local, smaller hospitals lose their skills and expertise due to lack of exposure or volume, this may negatively affect their reputation and perpetuate the cycle. Regionalization also reduces or eliminates competition between hospitals.[8]

Role of case-mix

When comparing hospitals with different volumes, the dissimilarity in crude health outcomes can be attributed to the disparity in illness severity, or case-mix.[9] It is crucial to determine if patients treated at the high and low volume centres are at equal risk for adverse outcomes. If one population is at higher risk for an adverse outcome, the difference in case-mix must be adjusted for in the analysis. For example, if a high volume hospital had a higher proportion of emergent or severely ill patients (i.e., high-risk), one would expect the high volume hospital to have poorer results than a low volume hospital that primarily treated elective cases or patients that were less ill. If the patient's severity or stage of illness was not adjusted for, one would unfairly conclude that the high volume hospital had disproportionately worse health outcomes. To obtain an unbiased estimate of volume's effect on outcome, all characteristics that could possibly influence the outcome should be controlled for using appropriate statistical

techniques. Methods for statistical adjustment of case mix vary in their complexity and their ability to capture different sources of case-mix differences.[10] The most basic method is to group patients by volume and stratify the outcomes by presence or absence of a risk factor. The second technique is to group patients by volume and account for different levels of risk by comparing the number of observed outcomes to the number of outcomes expected for the population with that level of risk. The third method is to develop a simple linear regression model with the hospital as the unit of observation and the outcome as the independent variable. A sophisticated method involves developing a multiple causal pathway to create simultaneous estimations of the effects of volume on outcome and the effect of outcome on volume. This approach also allows for the simultaneous estimation of the effects of physician and hospital volume on patient outcome.

Previous research

In 1979, Luft et al initiated research in this area when he and colleagues used discharge abstracts from a sample of American hospitals to examine whether a volume-outcome relationship existed for various surgical procedures.[11] A recent systematic review assessed 135 studies that encompassed 27 procedures and clinical conditions.[12] They found a strong association between increased volume and improved health outcome for AIDS treatment and surgeries in the following areas: pancreatic and esophageal cancer, abdominal aortic aneurysms, and pediatric cardiac conditions. They sighted poor methodological quality and insufficient adjustment for risk as challenges in this body of literature. A second review found that a volume-outcome relationship did not exist for peripheral vascular surgery when the studies undertook adequate risk adjustment.[13] They reviewed 37 studies that assessed volume at the physician and hospital level; most of the studies were retrospective and of poor methodological design.

To the best of our knowledge, there has not yet been a systematic review that specifically examines the volume-outcome relationship in pediatric health services.

Chapter 2 Objectives

The global objective of this project was to systematically identify and review studies that examined a volume-outcome relationship or evaluated the impact of regionalization of pediatric and youth health services.

Specifically, we aimed to:

1. Synthesize the available evidence on the impact of regionalization of pediatric services or studies that specifically examined a relationship between volume of provided service and pediatric health outcomes.
2. Identify subgroups of the pediatric population who may benefit from regionalized health services.
3. Aid in formulating policy to assist decision makers in the use of health service resources and identify potential areas for future research agendas.

Chapter 3 Methods

Identifying the literature

Comprehensive searches were developed by a medical research librarian (Appendix III). The following electronic databases were searched: Medline (1966-Oct 2002), EMBASE (1988-Oct 2002), HealthStar (1975-Sept 2002), and Cinhal (1982 to October 2002). PubMED was also searched in October 2002 to capture any recently published studies. Reference lists of the included studies were checked for potentially relevant articles. In addition, an author of each included study was contacted for additional studies. Finally, the Proceedings of the Society for Pediatric Research annual meetings were hand-searched for the years 2000-2002.

Selection of eligible studies

The identification of relevant studies involved two steps. The titles and, when available, abstracts, from the search output were independently screened by two reviewers. The full text of potentially relevant studies was obtained and using pre-determinedinclusion criteria, two reviewers independently assessed each study. Discrepancies were resolved by discussion between the two reviewers and, when required, consulting with a third party.

To be included in this systematic review, a study must have met the following five inclusion criteria. First, studies were considered for inclusion if they examined a volume-outcome relationship or evaluated the impact of regionalization of services. Studies that compared various levels of care were also included. Studies that reported volume and outcome data were excluded if the measurement of an association between volume and outcome was not an objective of the study or the authors did not draw conclusions about the volume-outcome relationship. Second, the study must have included a comparison group. The following study designs were considered eligible for inclusion: randomized controlled trial (RCT), clinical controlled trial, before/after, prospective or retrospective cohort, case control, or cross-sectional. Third, the participants had to be either: facilities or healthcare workers delivering pediatric health services, or any provider of child health (including governmental, non-governmental, or private organizations). Studies that included both adults and children were only included if children were analyzed separately. Fourth, the intervention of interest must have been any pediatric health service. Finally, the study must have measured at least one objective, quantified outcome, such as mortality, morbidity, length of stay (LOS), admission rates, cost of service, or user acceptability (such as travel costs or expressed satisfaction). Studies that used statistical models to simulate results were excluded.

Both published and unpublished research was considered for inclusion. Only English language studies were included, as excluding non-English studies has shown not to influence or bias the overall conclusions in meta-analyses of RCTs.[14]

Quality assessment

The methodological quality of the included studies was assessed using the partially validated Downs and Black checklist for randomized and non-randomized studies of health care interventions.[15] This tool measures quality in terms of reporting, external validity, internal validity (bias and confounding), and power. The highest possible score, referred to as quality index, is 29.

Because case-mix is critical in this area of study,[9 10] a second index of quality was used to assess the degree of adjustment for case-mix, or the extent of risk adjustment. Based on the thoroughness of risk adjustment, each study was assigned the appropriate case-mix level. Only variables that were adjusted for the pediatric volume-outcome relationship were included. Case-mix levels were modified from a previous NHS report and defined as:[16]

Case-mix Level	Criteria
0	No adjustment for case-mix
I	Adjusted for demographic variables
II	Adjusted for either co-morbidity or stage/severity of illness or better process of care, and may or may not be adjusted for demographic variables
III	Adjusted for demographic variables, co-morbidity, and stage or severity of illness

Data extraction

Data was extracted directly into an Excel database by one reviewer and checked for completeness and accuracy by a second. For each study, the following information was extracted: study design, objective, primary outcome, volume characteristics, patient characteristics, source of patient information, method of risk adjustment, and authors' conclusions. The levels of volume were characterized by a hospital's level of care, number of procedures conducted by a physician, number of beds in the compared hospitals, number of patients treated in each hospital, or years studied before and after the implementation of a regionalization program. Volume data was defined as categorical or continuous. Methods for including hospitals were recorded as random sampling, volunteering, a specific geographic location, database, or other methods. Extracted patient characteristics included the number, age, and risk score in each study group. Information about the condition and procedure was recorded, such as elective or emergency procedure and if the patients required a technical service or medical management. The source of patient information was classified as medical records, clinical database or registry, administrative database, or insurance forms. Collection of patient information was recorded as prospective or retrospective. Specifics concerning risk adjustment were extracted. In studies where a risk adjustment was performed, the variables used and the statistical methods to adjust for case-mix were recorded. Results

most relevant to the objective of the review were extracted; adjusted results were recorded where both adjusted and crude findings were reported. Lastly, we recorded whether the authors endorsed or opposed regionalization, concluded a threshold volume, and their overall conclusions.

Data analysis

Two reviewers performed a qualitative assessment and data was organized by clinical topics. Because many of the studies examined neonatal outcomes and the objectives of the studies were variable, the studies were divided into two categories: studies comparing the implementation of a regional program or hospitals with different volumes of deliveries and studies that compared hospitals with different levels of care (i.e.; different intensities of resources). Studies that included only babies being admitted to an NICU were separated from studies that included all births, regardless of health status. Within each clinical topic, the studies are presented such that the highest methodological quality studies are reported first; degree of adjustment for case-mix was considered a higher indicator than the Downs and Black quality index. Unless otherwise specified, measures of variance in the results table are 95% confidence intervals. A meta-analysis was not conducted because the included studies were too heterogeneous to pool. The volume definitions, outcomes examined, included patients, and degree of risk-adjustment varied too greatly within each topic.

Peer-review

Three individuals with expertise and an interest in this area were selected to form an external peer review committee. The peer-reviewers were asked to provide feedback on a draft report. The comments were collated and incorporated into the final report.

Chapter 4 Results

Results of literature searches

The searching of electronic databases generated 17,247 unique studies. The titles and, when available, the abstracts, were screened independently by two reviewers. The full manuscripts were obtained for 1,213 studies and were independently assessed for inclusion by two reviewers (Appendix IV). In total, 106 studies were included. From the electronic search output, 87 studies were deemed relevant. Eight more relevant studies were identified from the reference lists of the previously included studies. Two abstracts were found by hand-searching the Society of Pediatric Research (2000-2001) conference proceedings. The authors of included studies identified nine additional studies. A list of excluded studies is available from the authors upon request.

Study characteristics

The overall quality is presented in Table I and II. The median quality index of the studies was 21 (inter-quartile range 18,22) (Table I). The majority of the clinical areas had a median case-mix of 2 (table II).

Table I: Summary of Quality Indexes by Topic, Median (Interquartile Range)

Topic	Reporting	External Validity	Bias	Confounding	Power	Overall
Maximum score	11	3	7	6	2	29
Appendicitis (n=2)	9 (9,9)	3 (3,3)	5 (5,5)	4 (4,4)	1 (1,1)	22 (22,22)
Cancer (n=11)	9 (8,10)	3 (3,3)	5 (4,5)	4 (4,4)	1 (1,1)	21 (19.5,23)
Cardiac (n=12)	8.5 (8,10)	3 (0,3)	4 (4,5)	4.5 (4,5)	1 (1,1)	19.5 (17.25,22.25)
Level of Care (n=23)	9 (7.5,10)	3 (3,3)	5 (5,5)	4 (4,4)	1 (1,1)	21 (19.5,23)
Miscellaneous (n=5)	9 (7,10)	3 (3,3)	5 (5,5)	3 (3,3)	1 (1,1)	21 (18,22)
Neonatal (n=31)	8 (6,9)	3 (2.75,3)	5 (4,5)	3 (3,4)	1 (0,1)	20 (16,21.25)
NICU (n=8)	9.5 (9,10)	1.5 (0,3)	5 (4,5)	3.75 (3,4)	1 (0,1)	20.5 (18.5,21.25)
PICU (n=7)	10 (7,10.5)	3 (2,3)	5 (3.5,5)	3 (3,4)	1 (0,1)	21 (17.5,22)
Transplant (n=3)	10 (9,10)	3 (3,3)	5 (5,5)	4 (4,4)	1 (1,1)	23 (22,23)
Trauma	10	3 (3,3)	4 (4,5)	4 (3.5,4)	1 (0,1)	22 (19.5,22)

(n=7)	(8.5,10)

Table II: Summary of Level of Case-mix by Topic, Median (Interquartile Range)

Topic	Median (+/- IQ)
Appendicitis (n=2)	2 (2,2)
Cancer (n=11)	0 (0,1.5)
Cardiac (n=12)	2 (1.5,2)
Level of Care (n=23)	2 (1.5,2)
Miscellaneous (n=5)	0 (0,2)
Neonatal (n=31)	2 (1,2)
NICU (n=8)	2(2,2)
PICU (n=7)	2 (0,2)
Transplant (n=3)	1 (0.5,1.5)
Trauma (n=7)	2 (1,2)

The majority of the studies were conducted in the United States (n=63), followed by the United Kingdom (n=14), and Canada (n=7). The median year of publication was 1992 (inter-quartile range 1985, 1999). Administrative databases were the most common source of data (n=61), and the most common method for including patients was by geographical location (n=80).

Appendicitis

Two studies that were published in abstract form, investigated the association between length of stay and costs in hospitals with high versus low pediatric volume.[17,18] Both studies were carried out by the same group of investigators and involved a retrospective analysis of data obtained from state administrative databases. The studies had identical quality indexes of 22. In an effort to control for case-mix, the studies examined variables related to demographics and stage or severity of illness.

Chamberlain et al (2001) examined how length of hospital stay and costs for appendicitis were related to pediatric volume in California during 1995. High volume was defined as having a discharge rate of more than 10 children a day. When controlling for age, sex, race, and insurance status, complicated appendicitis treated in low volume hospitals was more cost effective and required a shorter hospital stay. For simple cases, there was no significant difference in length of stay, however the high volume hospitals still had higher costs.

The second study explored the relationship between length and cost of hospital stay in New York's high and low volume hospitals. Hospitals that discharged more than 2% of the state's total pediatric population were deemed high volume hospitals. Study

outcomes were adjusted for the following potential confounders: age, sex, race, Medicade insurance, resident of New York City, and presence of complicated appendicitis. Increased volume was significantly associated with shorter hospital stay and lower treatment costs.

Cancer

Eleven studies examined the impact of centralizing pediatric cancer treatment.[19-29]

Four studies examined how outcomes varied by hospital volume. Foreman (1996) explored possible explanations for the variation in survival of malignancies in the Southwest region of the United Kingdom. Between 1976-1985, children who were less than 15 years of age at diagnosis were retrospectively identified through pathology department records, the regional Pediatric Oncology Unit and tumor registries. Large treatment centres treated more than six patients a year and small centres treated between 1-5 children each year; this division was chosen because a previous study had shown that dividing at six patients resulted in a centre size effect. After controlling for the type of malignancy, Foreman found that survival in children with CNS tumors was significantly better in large centres for the first half of the study period only (1976-1980); there was no significant difference in survival based on hospital size for patients with leukemia or other cancers (QI=23). As part of a larger study investigating the impact of participation in Medical Research Council trials, Stiller et al (1989) examined whether centres with different volumes (defined as the number of patients treated each year) had different survival rates. They found that across the ten-year study period, rates of survival adjusted for age, were highest among the high volume centres and decreased with centre volume (QI=21). While the trend was consistent over time, the disparity in survival rates between centres with different volumes decreased over the study period. The authors did not test whether the differences between high, medium, and low volume centres were statistically significant. In 1999, Stiller et al reported survival rates in children less than 14 years of age when diagnosed with acute lymphoblastic leukemia (ALL) in relation to patterns of organization of care. Hospitals were defined as high (mean of at least six patients a year), medium (1-3 patients a year) or low volume (mean of less than one patient a year). Hospitals were further classified as being part of the UK Children's Cancer Study Group (UKCCSG) or non-UKCCSG hospitals that were not pediatric oncology centres. From 1980-1994, children were identified and data was retrospectively collected from the National Registry of Childhood Tumors. Stiller et al concluded that actuarial survival rates did not vary significantly between hospitals with different numbers of patients or between UKCCSG and other hospital. No adjustment for case-mix was made (QI=23). The final study was conducted by Danjoux et al (1996), who sought to explain discrepancies of five-year relapse-free survival of 108 Ontario children less than 16 years of age with posterior fossa medulloblastoma between 1977 and 1987. Children were identified by the Ontario Cancer Registry. By retrospectively comparing data obtained from discharge summaries, pathology reports, specialized cancer centre reports, and death certificates, the authors found that relapse-free survival was highest in large centres. They recommended that for relatively rare diseases that require complex treatment, such as medulloblastoma, a minimum number of cases should be treated to

ensure optimal expertise. The authors did not provide an estimation of the minimum number of cases, nor did they define large, medium, or small centre. The results were not controlled for any potential confounders (QI=22).

Six studies compared survival rates in children that were treated in different types of treatment centers. Using data from the National Registry of Childhood Tumors, Sanders et al (1988) retrospectively identified children under the age of fifteen that were diagnosed with retinoblastoma between 1969 and 1980 and were followed up for 17 years. The authors found that children had highest three-year survival rates when treated in London hospitals that jointly worked with the major UK referral centre than other specialist eye hospitals and survival was lowest in the remaining hospitals. After controlling for age and stage of cancer, Sanders et al (1994) found significantly better survival in London hospitals in children with unilateral tumors, however this relationship did not exist in children with bilateral tumors (QI=18). In 1988 Stiller published a study that compared survival rates in various treatment settings: pediatric oncology centres, teaching hospitals, and non-teaching hospitals. When controlling for age and year of diagnosis, four of the eight diagnostic groups (acute non-lymphoblastic leukemia, Non-Hodgkin's lymphoma, Ewing's tumor, and Rhabdomyosarcoma) showed significantly longer three-year survival rates when treated at a pediatric oncology centre. Meadows et al (1983) examined how place and type of treatment affected survival of children less than 15 years of age with acute lymphocytic leukemia (ALL) in Philadelphia and surrounding areas. The authors compared non-cancer to cancer centres; the latter had specialized medical support services, and a pediatric hematologist-oncologist planned care and therapy. Data was collected by the Greater Delaware Valley Pediatric Tumor Registry (1970-1975); results were not controlled for any potential confounding variables. The authors concluded that treatment at a cancer centre resulted in significantly greater survival (QI=22). Stiller et al (1994) conducted a study to identify possible prognostic factors for children less than 15 years of age with acute non-lymphocytic leukemia (ANLL). The National Registry of Childhood Tumors both retrospectively identified children with ANLL and supplied the data for analysis for 1971-1988. Survival rates since 1985 were compared between children treated in teaching hospitals versus those who received care in other hospitals. Survival rates were controlled for the following potential confounding variables: age, sex, morphological subgroups, white blood cell count, and the participation of children in trials. Stiller et al (1994) concluded that children had increased survival when treated in teaching hospitals, however this association disappeared when trial and non-trial children were analyzed separately (QI=19). Kramer et al (1984) described how management procedures and outcomes varied between cancer centres and non-cancer centres from 1970 to 1979 (QI=21). Cancer centres had a full-time pediatric hematologist/oncologist and non-cancer centres were not defined. Children from the Greater Delaware Valley were included if they were between 0-14 years of age and were diagnosed with one of the following types of cancer: Wilms' tumor, rhabdomyosarcoma, or medulloblastoma. Data was retrospectively obtained from the Greater Delaware Valley Pediatric Tumor Registry; no adjustment for case-mix was performed. For children suffering from rhabdomyosarcoma, three-year disease-free survival was greatest for those treated in cancer centres; this association was not observed in children with Wilms' tumor. There

was an increase in operative deaths in non-cancer centres for medulloblastoma in 1970-1974, but no association was found during the later part of the decade. In an effort to determine how survival differed by type of treatment centre, Duffner et al (1982) compared survival between children treated in university cancer centres to those receiving care in community hospitals (QI=23). Prior to analysis, university cancer centres and community hospitals were defined by the Chief of the Connecticut Tumor Registry. Using the Connecticut Tumor Registry, children with brain tumors were retrospectively identified from 1968-1979. A crude analysis of patient and computer records showed that a volume-outcome relationship appeared to be dependent on the type of brain tumor. For children with medulloblastomas and brainstem gliomas, survival was statistically better for children with treated in university cancer centres. There was no significant difference in survival for the following brain cancers: cerebellar or midline astrocytomas, astrocytomas (grades I, II, III, IV), ependymomas, or glioblastomas.

One study compared pediatric to general surgeons with regards to tumor removal and complication rate. Albright et al (2000) evaluated whether extent of brain tumor removal and/or frequency of complications varied between general or pediatric neurosurgeons. Using patients previously enrolled in three recent Children's Cancer Group studies, neurosurgical data sheets and operative reports of general and pediatric surgeons were compared. General surgeons generally had six months of pediatric neurosurgery training during their residency and pediatric neurologists were defined as either being a member of the American Society of Pediatric Neurosurgeons (ASPN) or being the designated pediatric neurosurgeon in their centre but not a member of ASPN. Pediatric neurosurgeons were more likely to extensively remove malignant pediatric brain tumors than general surgeons; the authors inferred that children with brain tumors should be operated on by pediatric neurosurgeons who perform numerous brain tumor operations per year. The results were not adjusted for case-mix (QI=20).

Cardiac

Twelve studies examined various aspects of the volume-outcome relationship in pediatric cardiac care.[8 30-40]

Four studies examined the impact of centralizing cardiac care in two areas: Sweden and New England states. Lundstrom et al (2000) compared 30-day mortality before and after centralization of pediatric open and closed cardiac surgery in Sweden (QI=19). In 1993, heart surgery in children under 18 years of age was centralized from 4 hospitals to the 2 hospitals with the lowest 30-day mortality. After controlling for severity of open-heart surgery, 30-day mortality rates decreased; no test of significance was conducted. The next three studies described and evaluated the achievements of the New England Regional Infant Cardiac Program (NERICP). Developed for the purpose of improving infant cardiac care, NERICP is a voluntary association of all hospitals that provide definitive care for infants with heart disease in the New England states. During 1969-1974, there was a consolidation of hospital services. Of the initial 11 hospitals in 1969, by 1974, there were five hospitals offering a full range of cardiac services, three offering limited surgical services, two no longer performing cardiac catheterization or surgery in

infants, and one that had discontinued all pediatric cardiology. The NERICP committee used data collected through state vital statistics and the NERICP database to calculate 30-day surgical mortality and mortality by first birthday for New England infants that required cardiac catheterization or cardiac surgery (QI=12). Mortality was adjusted for open versus closed operations. They found that 30-day surgical mortality decreased notably in all age groups, regardless of type of surgery, however the study suffered from many methodological errors. Flyer et al (1972) found that crude mortality was not significantly different in the eight small hospitals compared to three large hospitals (QI=20). However, since the development of NERICP, patients were being referred for treatment earlier and surgeries were being conducted at a younger age. Chameides et al (1975) compared mortality rates in the three large centres in New England (minimum 14 patients) to the seven small hospitals with 40 patients between 1968-1973 (QI=12). The authors found that crude mortality rates were not associated with centre volume; no risk adjustment was performed.

The volume-outcome relationship was specifically examined in seven studies. Hannan et al (1998) conducted the only study that considered the effect of both physician and hospital volume of congenital heart surgery on in-hospital mortality (QI=23). Annual cardiac volume of procedures performed in New York State were classified as high and low hospital volume (=100 cardiac surgeries versus <100) and high and low physician volume (=75 procedures versus <75). The breakpoints were chosen to yield the greatest difference in risk-adjusted mortality while preserving adequate volume in each group. Using retrospective data collected by New York's Cardiac Surgery Reporting System, risk-adjusted mortality rates were calculated for the four combinations of hospital and surgeon volumes from 1992-1995. The following factors were included in the risk-adjustment: age, "several clinical risk factors", and four procedure complexity categories. Hannan et al (1998) concluded that risk-adjusted in-hospital mortality was significantly reduced for both low and high complexity cases when surgeon and hospital volume were high. Using retrospective data from the Pediatric Cardiac Care Consortium registry, Scott et al (2001) examined 1,679 children who had ventricular septal defect (VSD) repair and 494 children undergoing arterial switch operation (ASO) between 1992-1996 (QI=22). After adjusting for a number of potential confounders (i.e., age, weight, presence of trisomy 21, associated risks, and duration of preoperative hospital stay), the authors concluded that increased hospital volume resulted in shorter length of stay for complex procedures (ASO), however this relationship was not evident for uncomplicated operations (VSD). Jenkins et al (1995) conducted a study to explore the relationship between volume of surgical repairs for pediatric congenital heart defects and in-hospital mortality (QI=23). Using retrospective statewide hospital discharge abstracts, Jenkins et al (1995) compared four categories of hospital volume in California (1988) and Massachusetts (1989). A hospital was considered high volume if they performed over 300 surgical repairs a year, 101-300 cases was classified as medium-high, 10-100 cases were low-medium, and hospitals performing less than 10 cases a year were low volume; volume thresholds were arbitrary. To control for severity of heart disease, each condition was classified into one of four case complexity categories. They concluded that the risk-adjusted mortality rate was significantly lower for children treated in hospitals with an annual volume above 300 cases each year. The fourth study by Sollano et al (1999) was

conducted in New York State's pediatric and adult population; only the methods and results pertaining to pediatric cardiac care are discussed (QI=23). The authors explored the relationship between hospital volume and in-hospital deaths for repair of congenital cardiac defects in children less than 18 years. Hospital volume was continuous and measured as hospital volume/100 cases. From 1990-1995, data was retrospectively collected by the Cardiac Advisory Committee. After adjusting for age and surgical complexity, Sollano et al (1999) concluded that pediatric patients undergoing cardiac surgical procedures had improved outcome as hospital volume increased. One study examined the effect of physician volume mortality in eleven surgeons in five selected United Kingdom hospitals (Stark 2000, QI=21). The range in surgeons' procedural volume was 19 to 235 over a one-year study period. Due to the large number of different types of operations and the resulting small data set, Stark et al (2000) were unable to conclude that physician volume was related to mortality. The authors stratified the results by type of cardiac procedure. In an effort to explain poor surgical results obtained by the Bristol Royal Infirmary, Spiegelhalter (2001) compared volume and 30-day mortality in England's 12 centres that performed repairs for congenital heart disease (QI=18). Volume was defined as the number of patients treated in each age group and the relative odds of death per additional patient per year was calculated. Data was collected from two retrospective sources (1991-1995): hospital episode statistics and voluntary submission to cardiac surgical register. Risk-adjustment was performed to control for open versus closed operations and type of procedure. Spiegelhalter (2001) concluded that success of open-heart surgery in children less than 1 year was correlated with increasing procedural volume, even when data from Bristol were excluded. Norwood et al (1988) explored how in-hospital and post-hospital mortality related to volume of arterial switch repairs in children older than 15 days (QI=17). Hospitals were divided into six volume categories: >50 repairs per year, 40-49, 30-39, 20-29, 10-19, or <10. The authors did not state how volume thresholds were determined. For the time period January 1, 1985 to June 1, 1987, data from twenty institutions collected through the Data Collection and Analysis Centre in Birmingham, Alabama was analyzed. Norwood et al (1988) concluded that mortality was inversely related to volume.

Chang et al (2000) attempted to identify factors, including case volume, that were associated with age at which the following surgical repairs were performed: atrial septal defect (ASD), ventricular septal defect (VSD), tetralogy of Fallot (TOF), and atrioventricular canal (AVC) (QI=19). The California Office of Statewide Health Planning and Development database was used to retrospectively identify Californian children less than 18 years who underwent repair for congenital heart disease repaired in 1995-1996. Volume was defined as the number of surgeries for a specific procedure. Age of definitive repair was adjusted for the type of congenital heart disease. Chang et al (2000) concluded that children were more likely to receive surgical repair at a younger age at centres with larger surgical volumes.

Neonatal: level of care

There were 23 studies that examined obstetric level of care from different perspectives: several studies compared outcomes for deliveries at basic versus sophisticated hospitals

or across different levels of care; some studies examined the effect of treating very low birth weight or high risk babies at different levels of care; two studies examined the effect of regionalized programs; and, one study evaluated utilization patterns across levels of care.[41-63]

Four studies compared deliveries at basic hospitals or centres to more technically sophisticated centres. Baruffi et al (1984) determined that basic hospitals were safe options for deliveries (QI=24). Using a matched sample of women delivering at either a basic medical centre or those delivering at a tertiary care hospital, Baruffi (1984) concluded there was no difference in neonatal mortality or morbidity; an extensive risk adjustment was performed. Sangala et al (1990) and Linder-Pelz et al (1990) collected data from medical records and minimally adjusted for risk difference. Linder-Pelz et al (1990) found that birth centres were a viable option for women with low obstetric risk singleton pregnancies, as there was no difference in obstetric outcomes between birth centres and conventional labour wards (QI=19). In a comparison of general practitioner versus consultant units, Clarke et al (1993) showed that perinatal mortality rates (adjusted for demographic characteristics) were substantially higher in the consultant units because of the selective referral of high-risk women from general practitioner to consultant units (QI=20). Conversely, Sangala et al (1990) concluded that the risk-adjusted perinatal mortality rate for infants weighing 2500 grams or more in isolated practitioner units was shockingly high compared to the consultant unit that cared for the high-risk deliveries (QI=19).

Two studies evaluated how outcomes differed for deliveries across levels of care. Both studies concluded that high-risk pregnancies were delivering in Level III institutions and therefore mortality rates were not statistically different between Level I and Level III hospitals. The studies supported regionalized delivery of services. The studies had quality indexes of 22 and 20, although Black et al (1984) (QI=20) did not adjust for case mix.

Thirteen studies concluded that low birth weight or high-risk babies had decreased risk of mortality when born in a level III, or tertiary care, hospital.[42-45 47 51 52 54-56 59 61 63] All but two studies used birth certificates,[44 52] or linked birth-death certificates[63]; Field et al (1991) collected data prospectively from medical records, while Verloove-Vanhorick et al (1994) used data from another study. The outcome in eleven studies was mortality[43 51 52 54 55 59 61 63] and in three studies, this was defined as death within twenty-eight days.[43 54 55] The birth weights for inclusion varied: three studies included births greater than 500 grams or one pound,[43 51 54] one study included births greater than 400 grams,[44] one study included infants with a birth weight greater than 1000 grams, one study included births less than 1500 grams, and the final study examined births over 2500 grams. One study only included babies that were at less than 28 weeks gestation and Bode et al (2001) examined infants born between 500-2000 grams. Three studies examined all births.[45 52 56] All studies controlled for risk factors, however three studies only adjusted for birth weight.[42 54 56] Overall the studies were of good quality with a median quality index of 22.

One study compared regionalized geographical areas to non-regionalized catchment areas and found no differences in outcomes (QI=16).[45] Rosenblatt et al (1986) determined that crude neonatal mortality rates of infants of at least 1000 grams in Washington State (regionalized) were not statistically different from those in Wales, a non-centralized area. Because of reporting differences in the two countries, the authors were unable to draw conclusions concerning very low birth weight infants. This study did not report risk-adjusted mortality rates.

Another study evaluated the effect of a regionalized program in Puerto Rico.[49] Becerra et al (1989) found a significant difference in mortality between the levels of care. They concluded that babies delivered in Level A facilities (similar to Level III) had increased neonatal mortality compared to babies delivered in Level C (comparable to Level I). The authors performed an extensive risk adjustment (QI=20).

Finally, a study by Lagoe et al (1986) examined the utilization patterns of high-risk babies in level III, II, and I hospitals (QI=18). High-risk babies were defined as those being extremely premature, premature with major problems, or neonates that died or were transferred to a different facility. Using data collected by New York's Statewide Planning and Research Corporation System and the State's Appropriateness Review of Neonatal Care Services, Lagoe et al (1986) concluded that length of stay decreased as the level of care increased for high-risk babies. The authors did not adjust for any potential confounding variables.

Neonatal: volume or regionalization

Twenty-eight studies examined the effect of regionalizing neonatal care or whether increased volume improves health outcome.[64-77, 65-80] This section is divided into five subgroups: recent studies, the Iowa experience, safety of small hospitals within a regional programme, volume-outcome relationship, and other miscellaneous objectives.

Because the methodological design in many of these studies is severely flawed, only those studies published after 1999 are described here (the details of all studies are provided in the appended tables). Three studies examined the effects of volume on health outcomes among neonates. The quality index ranged from 21-23. Two of the studies defined volume by the number of deliveries the hospital performed and this was determined from data collected retrospectively; the third study compared a regionalized country to a non-regionalized country. The International Neonatal Network, Scottish Neonatal Consultations, and Nurses Collaborative Study Group (2000) compared neonatal mortality at less than twenty-seven days in very low birth weight infants in Australia to those in Scotland. Australia has extensive centralization of neonatal care, whereas Scotland has tertiary and non-tertiary hospitals. After controlling for birth weight, the authors determined that neonatal outcome was significantly better in Australia. Moster et al (2001) determined if neonatal mortality in Norway (1967-1996) was predicted by the size of the maternity unit. Maternity units were grouped by the volume of annual births: less than 100, 101-500, 501-1000, 1001-2000, 2001-3000, and >3000. Using data collected by Statistics Norway (National Population Registry) and the

Medical Birth Registry of Norway, the following variables were controlled for: indicators of social class differences and urban-rural differences. Moster et al (2001) concluded that there was a significant decrease in risk-adjusted neonatal mortality as the size of the hospital increased, however this difference was small. Klein et al (2002) examined if a family doctor's annual volume of singleton deliveries affected maternal morbidity, 5-minute Apgar score, admission to a neonatal intensive care unit. Data was collected on family physicians practicing at British Columbia's Women's Hospital and Health Centre from April 1997- August 1998. High volume physicians delivered a minimum of 25 babies a year, medium volume physicians delivered between 12-24 babies, and low volume physicians delivered no more than 11 babies. Several potentially confounding variables were extracted from medical records: parity, ethnicity, lone parent status, age, pregnancy induced hypertension, gestation diabetes (type 1 and 2), head circumference at birth, gestational age, and birth weight. Klein et al (2002) found that adverse outcomes were not related to family physicians' delivery volume.

Hein et al conducted five studies that examined the impact of regionalization in Iowa (median QI=18).[68 69 81-83] Five studies examined results from the Iowa Statewide Perinatal Care Program for the following time periods: 1972 and 1982, 1972-1976, 1978-1979, 1982, and 1982-1983 and 1995-1996. All four studies concluded that regionalization resulted in a lowering of the neonatal mortality rate, however, two studies specified that it was the low birth weight infants born in tertiary care hospitals that benefit from regionalized care. The studies used administrative data collected from either Iowa State Health Department or Iowa Department of Public Health. Two of the studies failed to adjust for differences in case mix and two studies had substantial methodological limitations. In the fifth study, Hein also examined mortality rates across the three levels of care and concluded that small hospitals are capable of offering acceptable perinatal care providing they are part of a regionalized program.

Three studies described the performance of low volume hospitals in the context of a regionalized program.[70 74 84] Lumley (1988) explored the relationship between risk-adjusted perinatal outcome and hospital size in Victoria, Australia. Hospitals were classified into one of six groups: >2000 births, 1000-1999, 100-999, 50-99, 20-49, and <20 births. Both stillbirths and live births of at least 500 grams were included. Perinatal outcome was defined as death within 28 days and was adjusted for birth weight. Perinatal morbidity forms, linked perinatal death certificates and autopsy reports were used to conclude that small maternity units are safe within the framework of a regionalized program (QI=20). Fallis et al (1988) examined birth weight-specific perinatal mortality rates for small hospitals in Ontario, Newfoundland, and Saskatchewan. Birth-weight specific mortality rates were calculated using the number of early fetal deaths (>20 weeks gestation and weighing >500 grams) and early neonatal deaths (less than 7 days) per 1000 total births. Only infants weighing at least 2500 grams were included and results were adjusted for birth weight. Hospitals were classified as high (= 400 deliveries per year), medium (101-400), or low volume (<100) institutions. Using data from Hospital Medical Records Institute, the authors concluded that small hospital obstetrics does not result in increased risk-adjusted mortality for babies weighing more than 2500 grams (QI=16). Peddle et al (1983) examined how the Nova Scotia

Reproductive Care Program affected perinatal mortality and concluded that perinatal mortality rates in small hospitals were similar to those obtained in larger hospitals. The authors collected data (method not specified); they included all births of infants above 1000 grams and measured perinatal death within seven days. The authors did not adjust for any potential confounders (QI=16).

Three studies specifically examined the volume-outcome relationship. The studies used different outcome and volume measures. Williams (1979) measured volume on a continuous scale and concluded that mortality decreased as volume increased. This was a high quality study (QI=21), and the investigators adjusted for several demographic factors. LeFevre et al (1992) categorized volumes into four groups and determined that there was no relationship between physician volume and perinatal deaths. Deaths were controlled for birth weight, age, education, race, and smoking (QI=21). Campbell (1984) found that NICU admissions were inversely proportional to number of deliveries, however this study had severe methodological flaws (QI= 9), no risk adjustment was performed, and volume was not defined.

Nine studies examined the effects of implementing a regionalized neonatal program.[75 76 78 79 85-89] These studies were older (median publication year 1985) and were of variable methodological quality: quality ranged from 13 to 23, and only two studies controlled for variables other than birth weight.[85 87] No study controlled for co-morbidity. Rozier (1980) and Callon (1975) did not perform any sort of adjustment for case-mix and Hokestra et al (1981) stated the results were adjusted, but did not report what potential confounders were used in the analysis. One study performed a quasi-experimental study comparing eight funded regionalized areas to eight control regions.[86] Three studies concluded that establishing a regionalized neonatal program did reduce birth-weight adjusted neonatal mortality,[78 88 89] while a fourth study did a cost-benefit analysis and determined that regionalization was economically beneficial.[76] Shapiro (1982) believed that the decreasing mortality rate was related to regionalization, however the mortality rate in the control region also decreased. Four studies were unable to attribute regionalization to decreasing mortality or improved health outcomes[75 79 80 85]; one study did not perform any statistical test of significance.[75] Taylor et al (1989) examined the effect of closing a local hospital and concluded that neonatal mortality did not change.

Three studies compared additional aspects of regionalization. Bakketeig et al (1978) compared perinatal, fetal, and early neonatal mortality rates (less than seven days) in babies born in hospitals with different extent of obstetric aid.[90] After adjusting for parity, maternal age, marital status, and birth weight, Bakketeig et al (1978) concluded that the perinatal mortality rate was highest in babies delivered in counties with the lowest level of obstetric aid (QI=23). Siegel et al (1986) found no difference in one-year developmental outcomes for babies born in rural North Carolina, which is under the Rural Regional Perinatal Care program, compared to matched babies born in three North Carolina counties without regionalization (QI=22). Siegel et al (1986) adjusted developmental outcomes for maternal education, race, and birth weight. Deutchman et al (1985) compared the obstetric outcomes between obstetricians and family physicians.

Using the Goodwin index to adjust for demographic risks, the authors concluded that family physicians and obstetricians have comparable obstetric outcome (QI=18).

NICU

Eight studies examined various components of centralization to determine if there is a relationship between volume and outcome in the neonatal intensive care unit.[67,91-97]

The International Neonatal Network, Scottish Neonatal Consultants, Nurse Collaborative Study Group (2000) compared a volunteer sample of tertiary NICUs in regionalized Queensland, Australia to tertiary and non-tertiary NICUs in New South Wales, Scotland (QI=21). Using data prospectively collected from medical records, the authors determined the risk-adjusted mortality prior to hospital discharge for preterm (<31 weeks) or very low birth weight (<1500 g) infants. The CRIB score (comprised of birth weight, gestation, severity of congenital malformations, maximum and minimum appropriate fraction of inspired oxygen, and worst base deficit) was used to adjust for risk of death. The standardized mortality ratios were significantly higher in both non-tertiary and tertiary NICUs in Scotland.

Paneth et al (1982) examined New York City's neonatal mortality rates in three types of hospitals: Level III (14 regional centres that provided the entire spectrum of medical and surgical services for newborns), Level II (20 hospitals that cared for more complicated pregnancies and neonatal illness), and Level I (32 hospitals that cared solely for normal or mildly ill newborns). All singleton births between 501-2250 grams were identified from vital statistics for New York City between 1976-1978. After controlling for weight, gestational age, race, sex, and medical risk variables, the risk-adjusted neonatal mortality rate was significantly lower in Level III hospitals (QI=23).

Four studies examined whether increased volume in an NICU would result in improved neonatal outcome. Field et al (1999) compared survival of premature infants in large and small NICUs (QI=24). There were five large referral NICUs and twelve small units that were rarely sent referrals and provided NICU support for local catchment. Visiting nurses reviewed medical records prospectively and spoke to staff and observed each unit. CRIB was used to determine the number of risk-adjusted expected deaths. Field et al (1999) concluded that the actual deaths fell within the 95% confidence intervals of the CRIB expected deaths for small and large NICUs and that infants in the large centres were more ill. Fifty-four of the UK's 186 NICUs were randomly selected and volume of very low birth weight infants (<1500 grams) was defined as low (admitted <35 infants), medium (35-57 admissions), and high (>57 admissions) volume (UK Neonatal Group, 2002). Based on the unit's admission logs, death rates adjusted for case-mix did not differ between the three NICU volume levels. The following potentially confounding variables were controlled for: sex, birth weight, gestation, 5-minute Apgar, infant temperature on admission, and surfactant therapy, oxygen therapy, and blood gas measures within the first 12 hours after admission (QI=19). Horbar et al (1997) (QI=17) measured the association between several characteristics, including annual patient volume, and neonatal mortality rates of very low birth weight infants (501-1500 grams).

American NICUs that belonged to the Vermont Oxford Network were classified into quartiles based on number of admissions (<47, 47-77, 78-113, and >113). Data pertaining to infants from 501-1500 grams were obtained from the Vermont Oxford Network database. After adjusting for birth weight, race, gender, major birth defect, 1-minute Apgar score, mode of delivery, antenatal steroid use, small size for gestational age, location of birth, and prenatal care, the authors concluded that there was no association between patient volume and 28-day mortality or standardized neonatal mortality ratios. Horbar et al conducted a second study (1988) that examined eleven NICUs in the United States, Canada, and England. The included NICUs were respondents from a request for data from an international sample of neonatologists (QI=17). From 1983-1984, the number of live births between 701-1500 grams within the eleven NICUs ranged from 86 to 358. After controlling for birth weight, gender, and race, Horbar et al (1988) concluded that an institution's volume was not related to survival, with or without supplemental oxygen, at 28 days.

One study examined level of NICU care and NICU patient volume in California infants weighing 500-1800 grams (QI=20).[96] Level of care was defined as: Level III (regional or tertiary NICU that provided a full range of specialized neonatal care), Level II+ (expanded intermediate NICU that provided long-term ventilator support), Level II (intermediate NICU that cared for moderately sick infants), and Level I (no NICU). Level III and II+ centres were considered high volume if they admitted fifteen or more infants per day. Level II NICUs were defined as high volume if they admitted more than four patients a day. Using retrospective data from the birth-infant death cohort file and the California Office of Statewide Health Planning and Development, Phibbs et al (1996) concluded that increased volume in Level III NICUs improved infant outcome. Level III NICUs with low volume and Level II+ and II (regardless of volume) did not have significantly higher neonatal mortality than Level I hospitals that did not have an NICU. Mortality was defined as death within the first 28 days of life, or within the first year of life, if continuously hospitalized. Results were adjusted for weight, age, sex, race parental education and age, mother's HMO status, and source of insurance.

In a comparative study of the number of neonatologists or neonatal intensive care beds, Goodman et al (2002) assessed whether increasing availability of specialty services was associated with lower neonatal mortality (QI=22). Retrospective data from the National Centre for Health Statistics was collected on all births (>499 grams) and neonatal deaths (mortality within the first 27 days of life) that occurred in the United States during 1995. The number of neonatologists per 10,000 births was categorized as very high (11.6), high (7.5), medium (5.9), low (4.3), or very low (2.7). Likewise, the number of intensive care beds per 10,000 births was categorized as very high (59.3), high (40.7), medium (32.4), low (23.5), or very low (14.0). The following variables were used to adjust for case-mix: birth weight, sex, single versus multiple birth, maternal age, parity, race, education, martial status, and extent of prenatal care. The authors found no relationship between bed supply and risk of neonatal death. The risk was significantly lower in regions with a low supply of neonatologists compared to regions with a very low supply; there was no difference in survival between the other volume categories.

PICU

Seven studies examined different components of regionalization and its effect on the volume-outcome relationship in pediatric intensive care units (PICU).[98-104]

The objective of three of the studies was to determine if centralization of PICUs improved various health outcomes. Pearson et al (1997) compared PICU mortality rates in Trent, UK to those of Victoria, Australia in children between one month and 16 years of age (children younger than one month were included if they were admitted for a cardiac disorder). When adjusting independently for pediatric risk of mortality (PRISM) and pediatric index of mortality (PIM), the odds of dying in Trent, UK were double that of Victoria (QI=23). The authors did not report the source of their data, however the data was checked for accuracy by two investigators. The second study evaluated the impact of centralization of pediatric intensive care in Birmingham, UK (QI=21). Pearson et al (2001) examined the change in mortality rates and length of stay from 1991 to 1999 in children less than 15 years using routinely collected data from a variety of sources. While there were 34 fewer deaths per 100,000 children in the population following centralization, the difference was not statistically significant (p=0.05). The authors speculated that centralization of PICU might have contributed to the lower population mortality. In addition, length of PICU stay was significantly reduced by 29 hours and the rate of mechanical ventilation (proxy for severity of injury) before and after centralization was similar. The authors did not formally adjust for patient risk. The third study compared seven consecutive years of PICU admissions in a city with centralized care (Jacksonville, Florida) to national data. National data was collected from a 1988 survey of 96 PICUs in the US and data for Jacksonville was obtained from a computerized database for the years 1983-1989. The mortality rate of all PICU admissions in Jacksonville was 3.3% compared to the national rate of 8.8%. While DeNicola (1992) concluded that regionalized PICU was superior, there was no adjustment for case-mix and the study had substantial methodological flaws (QI=14).

In an effort to determine whether regionalization was required in Oregon and southwestern Washington, USA, Pollack et al (1991) compared mortality rates in tertiary PICUs (as defined by the Oregon Health Division) versus non-tertiary PICUs (QI=24). Between June 1 and Nov 30, 1986, post-graduate nurses collected data on children less than 18 years of age with head trauma or those requiring intubation for respiratory support. Risk scores were calculated using PRISM, age, and operative status at the end of the ICU admission day. In tertiary centres, the number of expected (predicted by PRISM) and observed deaths was identical. This was not the case for non-tertiary PICUs where there were significantly more deaths observed than expected. For severely injured children, the odds of dying in a non-tertiary PICU were eight times greater than if they were cared for in a tertiary PICU.

Three studies specifically examined whether a volume-outcome relationship existed for pediatric intensive care. To test the hypothesis that increasing volume reduces mortality, Tilford et al (2000) used data prospectively collected by the Pediatric Critical Care Study Group during 1993 (QI=21). Sixteen of twenty-one invited hospitals completed data

collection; centre volume ranged from 147-1,378 PICU admissions. Variables included in risk adjustment were age, PRISM, Pediatric Overall and Cerebral Performance Scores, age, university hospital affiliation, children's hospital, intermediate care unit, fellowship program, and number of PICU beds. Based on the finding that a 100-patient increase in a PICU volume resulted in a significant decrease in risk-adjusted mortality and length of stay, Tilford et al (2000) concluded that increasing volume reduced mortality in the PICU setting. The second study, conducted by Pollack et al (1994) explored several care factors, including volume, that were associated with mortality in a PICU (QI=21).[102] Volume was defined in terms of number of beds: PICUs with 4-6 beds were classified as low volume and PICUs with seven or more beds were considered to have high volume. Mortality was adjusted for age, PRISM, operative status, emergency or elective status, pre-ICU care area, transport method, pre-ICU cardiac massage, and system and cause of disease. Increased PICU size was not associated with decreasing mortality. The authors speculate that excluding PICUs with less than four beds may have prevented detecting a volume-outcome relationship. Lopez-Herce et al (2000) surveyed 34 PICUs in the Spanish public health system in order to describe pediatric intensive care in Spain. The authors categorized volume based on the distribution of beds reported in the questionnaires: >12, 7-12, and = 6. They found that mortality increased as the size of the ICU increased (mortality rate 4.8 in smaller units versus 6.3 in larger units). The authors did not attempt any sort of risk adjustment and the study was of low methodological quality (QI=11).

Transplant

Three studies examined the correlation between volume of transplant procedures and clinical outcome.[105-107] All three studies used data collected by clinical databases or registries. The three studies were methodologically strong with quality indices of 23 (Schurman 1999, Hosenpud 1994) and 22 (Cox 2001).

Schurman et al (1999) described the relationship between the number of kidney transplants conducted between 1987-1995 and graft survival. Hospitals that contributed data to the North American Pediatric Renal Transplant Cooperative Study (NAPRTCS) were classified into one of three volume classes: >100, 51-100, or =50 transplants each year. The authors controlled for a number of variables known to impact graft loss: donor source, recipient age, race, transplant year, HLS matching, prior transplantation, prior dialysis, transfusion history, and anti-T-cell antibody induction. They concluded that graft loss was significantly higher in low volume centres, however this relationship was no longer apparent when adjustment was made for anti-T-cell antibody induction.

Cox et al (2001) examined how mortality was related to number of pediatric liver transplants in American hospitals. Data were supplied by the Scientific Registry of the United Network of Organ Sharing (1994-1998) and hospitals were categorized as high (=15 transplants each year), medium-high (11-15), medium-low (6-10), or low (1-5) volume. The authors concluded that children less than two years of age had a significantly greater chance for survival when they were treated in centres that performed

at least ten liver transplants a year. For children older than two years of age, volume of liver transplants did not significantly affect mortality.

Hosenpud et al (1994) explored the effect of cardiac transplant volume on mortality in a pediatric and adult population. Hospitals were dichotomized into high (=9 transplants/year) or low volume institutions. Using data from United Network of Organ Sharing Scientific Registry, the authors demonstrated that high volume centres were associated with significantly less mortality between Oct 1987 and Dec 1991. The authors did not perform a risk adjustment for the pediatric transplant recipients.

Trauma

Seven studies examined various aspects of regionalization and volume-outcome relationship in trauma care settings.[108-114]

Two studies explored whether having an organized (regionalized) trauma system was beneficial with regard to reduced mortality and/or cost recovery. Cooper et al (1993) compared New York, a state without a centralized trauma system, to pediatric trauma centres that participate in the National Pediatric Trauma Registry (NPTR). An analysis of the discharge abstracts showed that patients that sought treatment for brain or internal injuries of moderately great severity (ISS 15-19) in pediatric trauma centres had survival rates 10 times greater than those treated in New York. All other types of injuries (skull, skeletal, open and closed wounds, vessels, crush, and cord/nerve) were examined across six ISS severity categories and the relationship between treatment at a pediatric trauma centre and decreased mortality did not exist. The authors did not perform tests of significance, instead they arbitrarily defined clinical significance as a twofold or greater difference in mortality rates greater than 10% or a fivefold or greater difference in mortality rates below 10%. Only injury severity score and site of injury were controlled for (QI=19). In a retrospective analysis of Chicago's medical and financial records before, during, and after the implementation of a pediatric trauma centre, Cosentino et al (1991) found that length of stay was longest one year post-regionalization, however mortality rates did not vary (QI=22). Length of stay decreased to pre-regionalization levels in years 2 and 3 post-regionalization. Total hospital charges were also higher after regionalization; collection rates remained unchanged. The authors did not adjust for any potentially confounding variables.

VanRooyen et al (1995) used data from the Chicago Health Department to explore what factors, including volume, were predictive of mortality in six Chicago urban pediatric trauma canters (QI=20). Although annual volume varied from 116-429 admissions, there was no difference in in-hospital mortality across the pediatric trauma centres. The authors did not control for case-mix.

Two studies specifically examined the relationship between case volume and health outcome. The second study to examine a volume-outcome relationship was conducted by Tepas et al (1998) (QI=22). They arbitrarily categorized hospitals as low (100-500 cases in the 5 year time period), medium (501-1000), or high volume (>1000). The Pediatric

Risk Indicator (a scoring system comprised of the Glasgow Coma Scale and absolute Pediatric Trauma Score) was used to adjust for patient risk. The authors found no relationship between case volume and mortality. In an effort to explore volume-outcome in pediatric and adult populations, Konovolina et al (1995) defined volume as the average number of patients seen in each hospital and defined physician volume as the mean annual number of patients per trauma centre (QI=18). By analysing retrospectively collected data provided by the Pennsylvania Trauma Outcome Study, they found a significant volume-outcome relationship only for adults with blunt injuries. No such relationship existed for pediatric patients or adults with penetrating injuries. The revised trauma score was used to adjust for case severity.

The final two studies compared pediatric mortality across various care settings in Pennsylvania. Data collected by the Pennsylvania Trauma Outcome Study allowed Potoka et al (2000) to determine if mortality varied when children were treated at pediatric trauma centres, adult trauma centres with qualifications to treat children, level II adult trauma centres, or Level I adult trauma centres (QI=23). After adjusting for demographic variables and trauma severity, Potoka et al (2000) concluded that children have significantly less mortality when treated at a pediatric trauma centre or an adult trauma centre with pediatric qualifications and that severely injured children faired best at pediatric trauma centres. Nakayama et al (1992) evaluated mortality in pediatric trauma centres versus urban or rural non-pediatric trauma centres (QI=22). With data supplied by the Pediatric Regional Resource Trauma Centres, the authors used the TRISS method that calculates risk-adjusted probability of survival. Analyses stratified by probability of survival, based on TRISS methodology, showed no significant differences between centres. When probability of survival was between 0.3 and 0.6, there was a trend for increased survival in pediatric trauma centres, however this difference did not reach statistical significance.

Miscellaneous

Five studies examined the volume-outcome relationship in different clinical areas.[115-119]

Two studies examined the role of specialist versus non-specialist care. Baumer et al (1997) compared children with diabetes who were treated by a specialist with a caseload greater than 40 to those treated by non-specialists with a caseload less than 40 children (QI=23). In the South Western region of England, members of the diabetes management team completed a questionnaire regarding each child's care and required services. Results were adjusted for the following potential confounders: age, living arrangements, duration of diabetes, family history of diabetes, Townsend deprivation index, insulin dose, and number of injections per day. The authors found that when compared to those treated by non-specialists with a lower case load, children cared for by specialists had better clinical outcomes, including lower glycated haemoglobin, least time in hospital at diagnosis, increased likelihood of having weight and height plotted, microalbuminuria and retinopathy screening, and fewer dietician consultations. The second study (Snow et al 1996) compared cost-effectiveness and quality of care between pediatric and general urologists for ureteroneocustostomy in children (QI=21). Medical records from Primary

Children's Medical Centre at the University of Utah School of Medicine were reviewed from 1990-1993. Snow et al (1996) concluded that pediatric urologists were more cost-effective, as mean hospital charges and length of hospital stay were significantly less compared to children operated on by a general urologist. Although pediatric urologists had fewer complications, the difference was not statistically significant; no risk adjustment was performed.

Bratton et al (2001) evaluated whether hospital pediatric caseload was associated with surgical management of intussusception (QI=22). Based on number of pediatric admission in 1996, hospitals were divided into three categories: 0-3000, 3001-10000, or >10 000. Using computerized information that was available for all patients hospitalized in the state of Washington from 1987-1996, Bratton et al (2001) concluded that children treated in high volume hospitals had less surgical intervention, decreased length of stay, and lower hospital charges. This relationship remained true after adjusting for age and presence of co-existing conditions.

Two studies explored a relationship between surgeons' experience and clinical outcome. Albright et al (1999) compared neurological outcomes between three surgeons with different years of experience performing tethered cord release (QI=17). One hundred consecutive patients requiring tethered cord release between March 1996 and May 1988 were analyzed; data was extracted from medical records. The authors did not explain their rationale for the selection of neurosurgeons. The frequency of postoperative neurological deficit could not be explained by the neurosurgeons' experience. The authors did not adjust for differences in case-mix and the authors state that outcome analysis is complex. Witt et al (1998) explored the concept of an operative learning curve for surgeons performing palatoplasty (QI=16). Children who underwent a cleft palate repair between 1978 to1990 were identified through a computerized data system. It was not clear where the children where being treated or the source of the database, and no assessment for case-mix was performed. The authors showed a relationship between the number of procedures performed and the success rate (no need for secondary palatal management) for one of the nine surgeons included in the study. There was no relationship observed between number of procedures and success rates for the other eight surgeons due to the small number of procedures each one performed. The authors commented that the number of cases needed to develop proficiency is difficult to quantify.

Chapter 5 Discussion

Findings

Overall

We reviewed 106 studies that examined volume-outcome relationship or the impact of regionalization on pediatric health services. The studies involved many different clinical areas; used different volume definitions and patient inc lusion criteria; and examined a variety of outcomes. Overall, the evidence demonstrates that high-risk patients fare better in a regionalized system. In general, low birth weight infants or high-risk pregnancies have improved outcomes when treated at high volume and/or level institutions. Also, critically ill children or those severely injured tend to have improved health outcomes when admitted to a high volume and/or level hospital. Children undergoing surgical procedures have improved outcomes when treated in a high volume institution. The findings endorse a volume-outcome relationship fairly consistently, but they do appear to be dependent on the procedures or specialties.

Clinical areas and role of risk adjustment

Appendicitis

It is unclear whe ther a volume -outcome relationship exists for the treatment of appendicitis as the two studies had conflicting results. Both studies, performed by the same group of researchers, adjusted for demographic risk factors and stage/severity of illness and had identical methodological quality. Numerous adult surgical procedures have displayed a volume-outcome relationship and other pediatric surgeries included in this review discovered a relationship between increasing volume and improved health outcome.

Cancer

Treatment at cancer centres or university/teaching hospitals and by specialists with a higher caseload appears to improve survival, although the studies performed minimal risk-adjustment. This relationship also appears to be specific to certain types of cancers. Of the four studies that examined the effect of volume on improved health outcomes, increased volume improved outcomes for CNS tumors and medulloblastomas, and the relationship was unclear for acute lympohpmatic leukemia. Six studies looked at the survival in different types of treatment centres; the treatment centres were diverse and often poorly defined and this makes it difficult to make overall conclusions regarding the optimal place for cancer treatment. The results suggest that "cancer centres" do have improved survival, as all of the studies showed increased survival for for at least one type of cancer they examined. Again, this relationship between adverse health outcome and increased volume is consistent with findings from the adult cancer literature.[120] It is unclear if increased volume is the decisive de terminant of increased survival or whether improved outcomes are the result of other factors associated with treatment at a more sophisticated centre. The three studies had minimal or no adjustment for potential

confounders, therefore it cannot be ruled out that the observed relationship is not caused by specialists or cancer centres treating children who are less ill. If children under the care of a specialist or receiving care at a more sophisticated treatment centre were diagnosed and treated earlier in the course of their illness compared to children at a low volume, non-cancer centre, a spurious relationship between volume and outcome could be present.

Cardiac

Regionalization of cardiac procedures and increasing procedural volume is positively correlated to increased survival, especially for complicated surgeries and high-risk infants. This relationship is consistent with that seen in adult cardiac surgical procedures.[121] Two studies did not conclude that increased volume resulted in improved outcomes. Potential explanations for this finding are poor methodological quality and lack of statistical power. The study by Chameides et al had substantial methodological flaws. The authors did not perform any statistical tests of association, but concluded that volume was not predictive of outcome. Secondly, volume was inadequately defined and there was no risk adjustment. It would be more interesting to determine how, since the implementation of the cardiac care program, risk-adjusted mortality rates compare across hospitals with different volumes. If the regionalization program was functioning optimally, high volume institutions should have similar risk-adjusted outcomes. Stark et al also did not detect a volume-outcome relationship, but attributed this to small sample size caused by subgrouping of patients into numerous procedure categories.

Neonatal: level of care

There were several different themes that emerged from the 23 studies that examined outcome by level of care. Thirteen studies consistently demonstrated that low-birth weight babies had improved outcomes when treated at higher-level institutions. These studies were generally of high methodological quality and most adjusted for severity of illness. Four studies compared basic birthing centres or hospitals with more sophisticated or technical birthing units. Three studies with high level of case-mix adjustment demonstrated that there was no difference in outcome; one study commented that the lack of difference was due to the effectiveness of the regionalized system that resulted in the transfer of high-risk deliveries to the more technical centres. The one study that concluded differently was of poorer quality and did not adjust as thoroughly for differences in case mix. One study found that there was no difference in outcomes between two geographic areas, one with regionalized delivery of care and one without. However, the authors were unable to examine outcomes for very low birth weight infants due to reporting differences in the two countries and the authors did not adjust for case mix differences. Another study examined utilization patterns of high-risk babies among different levels of care and found that length of stay decreased as the level of care increased, although the authors did not adjust for any potential confounders. Finally, one study evaluated the effect of a recently implemented regionalized sys tem in Puerto Rico and found that the primary and secondary hospitals had higher neonatal deaths than expected. This study highlights the importance of designing programs to meet the

specific needs of different regions and evaluating programs post-implementation to ensure that they are having the intended effect.

Neonatal: volume or regionalization

Numerous studies examined how regionalization affected health outcomes of newborns. The strongest conclusion that can be derived is regarding the safety of small hospitals operating within a regionalized area. Four studies compared mortality in high and low volume hospitals to a geographic area with an existing regionalized program; each study concluded that low volume hospitals had mortality rates similar to their larger counterparts. Three of the four studies adjusted for risk. Seventeen studies examined the effects of implementing thirteen distinct regionalized perinatal programs and the evidence is inconclusive. Five examined perinatal rates in Iowa and concluded that Iowa's program has improved neonatal mortality; three other programs had similar results. All but one of these studies controlled for case-mix. One study found regionalization to be cost-effective. Four studies were unable to conclude that implementing a regionalized program resulted in decreased mortality, and once again, only one study failed to adjust for patient risk. There were three studies that compared regionalized areas to non-regionalized areas. Neonates born in a country with regionalized care (Australia) had lower mortality than those born in Scotland, a country that lacked regionalized perinatal care. Two studies compared counties in North Carolina that received funding to develop a regionalized program, to counties without such a program. Although mortality rates did decrease over the thirteen years studied, the authors were unable to attribute the decrease to the regionalized program. Also, the program did not show differences in long-term outcomes, in terms of neurological and physical development, in babies born in a regionalized area versus those born in the non-regionalized control region.

No definitive conclusions can be drawn about how neonatal volume affects mortality. Five studies compared neonatal mortality in hospitals with various volumes. Three studies concluded that increased volume results in improved survival and two of the three studies adjusted for case-mix. However, two studies determined that outcome did not differ in hospitals with different volumes; both studies controlled for differences in patient risk.

In summary, low volume hospitals in the context of a regionalized system do not show an increased risk of mortality. It is unclear if the implementation of a regionalized program or increased number of deliveries by a hospital is beneficial in decreasing mortality. Approximately half of the before-after study designs did not show decreased mortality after the implementation of a regionalized program. This finding may have been due to the authors not controlling for pre-existing decreases in perinatal mortality rates. Also, perinatal mortality has been decreasing over time and it can be difficult to attribute any change in mortality rates to the regionalization program.

NICU

In the NICU setting, increased level of care appears to result in improved survival more so than increased volume. Four studies assessed mortality in high and low volume NICUs and none found an inverse relationship between decreased mortality and increased volume. However, there is more conclusive evidence in regards to the effect of resource availability. The International Neonatal Network concluded that, when compared to Australia's regionalized NICU system, Scotland's non-centralized NICUs were responsible for unnecessary deaths. Paneth et al concluded that infants had improved survival in Level III hospitals and that this was due to better techniques and more extensive resources. Two studies looked at both volume and resource availability. Phibbs et al deduced that mortality decreased in high volume Level III institutions, and mortality between low volume Level III and Level II hospitals did not differ. In contrast Goodman et al examined the effect of beds (proxy for size of hospital) and neonataologist availability and found that infants had increased survival when there was a higher number of neonatologists. All studies incorporated risk-adjustment for at least demographic variables.

PICU

The PICU studies showed a similar pattern as seen in NICU. Three studies examined the effects of implementing regionalized programs over time or compared a regionalized geographical area to a non-regionalized area. Pearson (1997) concluded that Trent's non-regionalized PICU system was accountable for deaths that would have been prevented in Australia's regionalized PICU. DeNicola concluded that regionalization of PICUs in Jacksonville resulted in lower pediatric mortality compared to the national average. This study had serious methodological flaws and did not perform any risk adjustment or statistical tests of association. In comparison, a second higher quality study by Pearson (2001) failed to find a difference in mortality, but the mortality rate in the overall population decreased and length of stay in PICU was significantly reduced. Overall, regionalization of PICU appears to be beneficial.

It is uncertain whether there is a relationship between PICU volume and outcome. Tilford found an inverse relationship between volume and outcome, however Pollack (1994) was unable to find such an association. Pollack limited the low volume category to at least four beds, and suggested that the lack of extremely low volume prevented finding a significant volume-outcome relationship. A third study also concluded that volume was not related to improved outcome, however no risk adjustment was performed. Thus, the increased mortality at high volume PICUs may be due to more severe patients being treated at high volume institutions.

Finally, one study examined outcome for different levels of care. Pollack (1991) concluded that the observed deaths equaled that of the PRISM predicted deaths in Level III PICUs, however there was excess death in non-tertiary PICUs.

Transplant

Three studies evaluated volume-outcome relationships for different types of transplants; the studies varied in the extent of risk-adjustment. Hosepud et al concluded that children receiving heart transplants had significantly better survival when the centre volume was greater than nine. The decision to dichotomize volume into less than nine and nine or above was made post-hoc and no risk adjustment was performed. Cox et al determined that increasing number of liver transplants was positively correlated to increased survival, however this relationship was only existed for children less than two years of age. Schurman et al found that graft survival of kidney transplants was positively related to increased volume. However, the correlation was no longer significant once the findings were adjusted for anti-T-cell antibody induction. These results illustrate how controlling for potential confounding variables (age in Cox's study and anti-T-cell antibody induction in Schurman's study) can change the results.

Trauma

Seven studies examined regionalization in a trauma setting and the results were inconclusive. Two studies attempted to determine if a regionalized trauma centre is advantageous from two perspectives: effect on mortality and economic effectiveness. In a study examining pediatric and adult populations, Cooper et al concluded that pediatric mortality in non-regionalized New York was not statistically different than in centres that participated in the National Pediatric Trauma Registry. Mortality was adjusted for severity and location, and the study was of moderate methodological quality (QI=19). Costenino also concluded that regionalization of trauma centres did not reduce mortality rates and total hospital charges increased, however the authors did not control for case mix (QI=22). Based on these two studies, regionalization of trauma centres does not appear to improve health outcome or to be more cost effective.

It is unclear if the hospital's extent of specialization of pediatric care or pediatric trauma designation is related to volume. Two studies compared pediatric trauma centres to various other institutions and neither study was able to show the superiority of the pediatric trauma centre. Potoka found mortality rates to be similar in pediatric and adult trauma centres with qualifications to treat pediatric patients; mortality was significantly higher in adult centres that lacked pediatric qualifications. When contrasting pediatric trauma centres to urban and rural centres, Nakayama found no difference in mortality. The studies were of similar methodological quality (22) and degree of risk adjustment. Similar to NICU and PICU reports, volume of patients does not seem to influence or predict mortality. Three studies with a breadth of quality indexes and degree of risk-adjustment concluded that increased volume does not result in a decreased mortality rate (Van Rooen, Konovolina, Tepas).

Miscellaneous

Five studies were divided into two categories: care by specialist versus non-specialists and the effect of increasing volume on health outcome. The two studies that explored the effect of specialist versus non-specialist care found that children treated by a specialist had better clinical outcomes (Baumer), and shorter length of hospital stay and more cost-

effective treatment (Snow). Baumer adjusted for patient risk, whereas Snow did not. The same pattern was seen in two of the studies examining the volume-outcome relationship (Bratton and Witt). Both studies concluded that increased volume produced advantageous outcomes: decreased length of stay and cost-effective treatment (Bratton), and less need for secondary palatal management (Witt). Witt did not account for any patient risk variables, while Bratton controlled for age and co-morbidities. The third study (Albright) was not able to correlate postoperative neurological deficit and the number of years of experience in three surgeons. The authors did not state how they chose the three surgeons, nor did they control for case mix.

Quality of studies and limitations

Overall quality

The overall quality index of the included studies was 21 (intra-quartile range 18, 22). Year of publication and quality index of the study were significantly correlated ($r=0.35$, $p<0.01$). The pattern of increasing methodological quality was the most extreme in areas that began volume-outcome research long ago. The first studies conducted examining the volume-outcome relationship were in cardiac and neonatal settings (1972 and 1975, respectively). For these topics, there was a significant positive relationship between methodological quality and year of publication (neonatal $r=0.41$, $p<0.02$; cardiac $r=0.60$, $p<0.05$). Improvement of quality over time is consistent with previous literature, although areas remain where quality could still be improved.[122]

The two sections with the highest median methodological quality were trauma [22 (19.5, 22)] and PICU [21 (17.5, 22)]. The appendicitis, transplants, and miscellaneous sections also had high quality indexes, however these included four or less studies. In studies examining regionalization of PICU, three of the four risk-adjusted outcomes showed that there was a significant decrease in mortality in higher volume institutions. The three studies that did not adjust for case-mix found variable results concerning the volume-outcome relationship, and one speculated that mortality in higher volume hospitals was higher due to more severe patients being seen at high volume institutions. Two studies did not adjust for case-mix and both concluded that increased volume or centralization of trauma care did not improve health outcomes. Four studies performed a risk assessment; these concluded that regionalization decreased mortality rates and that children with some types of injuries fared better in pediatric trauma centres or adult centres that were qualified to treat children. The three other studies did not demonstrate a volume-outcome relationship. A similar pattern is seen within the trauma studies.

Observational studies

The nature of this study question does not lend itself to the gold standard study design: the randomized controlled trial. Because it is not feasible, nor ethical, to address the volume-outcome relationship in a randomized controlled trial, there is inherent bias associated with this body of literature.[6]

Compared to experimental designs, observational studies are more prone to confounding.[123] If confounding is present, factors other than differences in volume may explain, or partially explain, the volume-outcome relationship. This is why it is imperative to adjust for case-mix and other potential confounders, in order to eliminate the possibility that disparities in patient illness severity or other variables explain the relationship. For example, if the most severe patients were treated in low volume hospitals, then it would be expected that those institutions would have a higher mortality rate than high volume hospitals that treated a greater proportion of less severe patients. If one only compares crude mortality rates, then the conclusion would be confounded by the differences in patient severity.

A further potential source of bias is timing of the study. In prospective studies, the researchers have more control over what variables are collected and how the data is collected, whereas retrospective studies are more susceptible to measurement errors.[123,124] Of the 106 studies in this review, only nine studied patients prospectively. Many of the included studies relied on routinely collected sources of data. While the use of secondary data sources can be more efficient, the investigators have less control over the quality of the data. Also, the opportunity to control for confounding is limited to the variables in the database; this may not coincide with suspected confounders.

Finally, before/after study designs have an increased risk of bias. Due to advancements in surgical technique, pre and post-operative treatment, and anesthesiology, mortality and complication rates have decreased overtime.[125] It is possible that studies that concluded increased volume was associated with improved outcome may be explained by other factors, such as enhanced treatment of such patients.

Limitations with the literature

Outcomes measured

Outcomes were not measured consistently among the included studies. Although the majority of the studies did provide some measure of mortality, there was still variation; this problem was particularly evident in the neonatal literature. The time points used to measure neonatal mortality differed in North America and Europe. Traditionally, North Americans defined neonatal mortality as death within 28 days, where Europeans only measured up to seven days. This is especially problematic in studies comparing differences in neonatal mortality between a European and North American country (e.g., Wales vs. Washington). A second inconsistency was the type of births included in the mortality measurement. The studies used varying combinations of fetal deaths, stillbirths, live births, neonatal births, perinatal births, singleton births, and multiple births.

Studies examining the volume outcome relationship in surgical procedures tended to measure 30-day mortality. It is possible that relatively short period of follow-up is insufficient to detect differences in outcome or that differences in outcome may be reflective of the hospital's discharge policies and not patient outcome.[16] If a hospital has a more lenient discharge policy, their mortality rates would appear lower because the

subject is not a patient in the hospital and, thus a death would not be observed.[10] Also, mortality is an inappropriate choice of outcome when evaluating low risk procedures.[16]

A similar issue is apparent in before-after studies. Long-term evaluation is required to evaluate the true impact of a regionalized program. Often dramatic changes may be seen immediately following the implementation of a program, however long-term follow-up is necessary in order to allow patterns to stabilize and to evaluate the program's full effect. In addition, changes seen post-implementation may simply reflect the continuation of a pre-existing downward trend (e.g., decreasing neonatal mortality rates) that is independent of the effect of the program. Three studies examined the implementation of a regionalized program and measured outcomes one year after the program began; these concluded that regionalization improved patient outcomes.[32 87 89] A more accurate testament of the programs' effects would be seen several years after its existence.

Objectives studied

Given the array of objectives studied within the included studies it was not possible to conclude that a specific number of studies reported a positive relationship between increased volume and improved outcome. While the aim of many studies was to find an association between volume and outcome, some studies did not expect to find a difference in patient volume and outcome. For example, when the objective of the study was to determine if regionalization was effective, there should not be a difference in risk-adjusted mortality. In a regionalized system, the patients with an increased risk of dying should be treated at high volume/level institutions, and after adjustment for case-mix, the mortality rates should be similar across the region. In this case, one would not expect a volume-outcome relationship.

Definition of volume

There were several examples where the authors defined volume after the data was collected. When volume is defined post hoc there is the danger that the choice of thresholds is biased. In three examples, the authors specifically chose to differentiate between high and low volume because it resulted in the greatest difference in outcomes.[33 106 115] In twenty studies, the investigators provided a rationale for their volume cut-offs.[24 25 39 42 43 45 46 50 52 54 56 57 61 70 72 74 96 103 111 112] The findings from these studies may be less biased.

Hospital versus patient volume

Previous research has shown that hospital and physician volume are significantly related to outcome.[126] Hughes et al (1987) found that increased hospital volume and lower proportion of patients operated on by low-volume surgeons resulted in significantly improved patient outcomes for most procedures they examined. In only six instances did the studies define volume as the number of procedures the physician performed.[72 73 92 115 117 119] Three studies examined the number of deliveries by physician and only one concluded that increasing number of deliveries did not improve neonatal morality or

morbidity.[73] Interestingly, the two studies that found an inverse relationship between decreased mortality and increased volume performed a more exhaustive risk adjustment than Klein et al (2002). The four other studies examined physician volume for ureteroneocystostomy, NICU care, palatoplasty, and diabetes treatment. All four studies controlled for potentially confounding variables. Because of the great collinearity between hospital and physician volume, it has yet to be determined whether hospital volume or physician volume is a greater predictor of health outcome.

Power

Many of the studies used data collected by administrative or clinical databases. Thus we assumed that studies with a population-based sample that found significant differences were adequately powered to find such a difference. However, when studies did not use a large database, in only two instances did the authors include a sample size calculation in their methodology.[97,127] For such studies, it is not clear if there truly was no association between volume and outcome or if the study was underpowered to find such an association.

Chapter 6 Conclusions and recommendations

Implications for delivery of health services

There were several themes that emerged from the wealth of information reviewed:

1) Outcomes for surgical procedures were generally more favorable in high volume hospitals. This trend was observed in cardiac operations, malignant tumor removal, and transplants procedures; however it was unclear for the treatment of appendicitis. Thus for some technically demanding procedures, volume does appear to be inversely related to improved health outcome.

2) Among high-risk children, resource availability appears to be a larger predictor of health outcome than volume. The evidence suggests that severely ill children fare better in level III institutions. Patients admitted to NICU, PICU, and trauma centers tended to have more favorable health outcomes in institutions with more resources.

3) Outcomes for high-risk pregnancies and premature babies were improved when delivered in Level III hospitals. Thus, it is important that expecting mothers are adequately screened so high-risk deliveries can be identified and delivered in a hospital with the appropriate level of care.

4) Another component that is critical to optimizing delivery of health services is the existence of an effective transportation network so patients can receive treatment at the most suitable institution.

5) Studies that examined outcomes of pediatric specialists versus general surgeons consistently showed that treatment by pediatric physicians resulted in improved outcomes and/or was more cost-effective. This finding supports the position that children are not merely "little adults" and that relying on adult based studies and extrapolating the results to children may not be in the best interest of child health.[128] Therefore, it is our recommendation that children be treated at hospitals and by physicians that coincides with their individual health risks.

Implications for research

The review also provided direction for future research in this area. This falls into two main categories: 1) how to improve on methodological quality of research; and, 2) questions that remain to be answered or clarified.

1) Methodological Quality:

a. Because of the biases inherent to the design of studies in this area, it is imperative that appropriate methods are used to control for confounding. Researchers need to ensure adequate adjustment for case-mix and other potential confounders.

b. In order to check accuracy and completeness, data should be collected in a prospective manner. If prospective data collection is not feasible, data obtained from clinical registries are superior to administrative databases.[6]

c. Comparability of studies would be enhanced if studies consistently measured the same outcomes and used similar study designs. The vast heterogeneity within the included studies eliminated the option of performing a meta-analysis.

d. Long-term studies needed to fully evaluate long-term impact of regionalized programs.

2) Research Questions:

a. There was only one study that examined the effect of physician and hospital volume on health outcomes. There is a need to differentiate the effects of physician and hospital volume, as this may affect the optimal delivery of health services.

b. It would be helpful (to policy and decision makers alike) to conduct a similar studies on pediatric volume-outcome relationships within more recently established pediatric health service delivery settings such as: "children's hospitals within hospitals"; children's hospital within a regional health system, maternal child settings and/or other such mergers (e.g. pediatric/rehab etc. etc), community care access rehabilitation centres etc.

c. Very few studies concluded a threshold volume for the minimum number of procedures required that would result in an "acceptable" number of poor outcomes. A useful research approach that could aid in policy development would be to develop statistical models for the clinical area under question to explore the effect of transferring patients from low volume to high volume hospitals. Such studies would be able to determine the change in outcome for both high-risk patients and if all patients obtained health services at high volume hospitals. Change et al (2002) developed a theoretical model that showed the change in mortality when patients in low- and medium-volume were referred to high volume institutions.[129]

d. Further research is not recommended in the area of neonatal/level of care. It is well established that high-risk and low birth weigh babies have an increased chance of survival when born in level III hospitals.

References

1. Dimick JB, Stanley JC, Axelrod DA *et al*. Variation in death rate after abdominal aortic aneurysmectomy in the United States: impact of hospital volume, gender and age. Ann Surg 2002; 235(4):579-85.

2. Hannan EL, Racz M, Ryan TJ *et al*. Coronary angioplasty volume-outcome relationships for hospitals and cardiologists. JAMA 1997; 277(11):892-8.

3. Smyth RL, Weindling AL. Research in children: ethical and scientific aspects . Lancet 354(suppl II):S1121-4.

4. Tilford JM, Simpson PM, Green JW, Lensing S, Fiser DH. Volume-outcome relationships in pediatric intensive care units. 2000; 106((2 Pt 1)):289-94.

5. Thompson DR, Clemmer TP, Applefeld JJ *et al*. Regionalization of critical care medicine: task force report of the American College of Critical Care Medicine. Critical Care Medicine 1994; 22:1306-13.

6. Luft HS, Garnick DW, Mark DH, McPhee SJ. Hospital volume, physician volume, and patient outcomes: assessing the evidence. Ann Arbor Michegan: Health Administration Press Perspectives, 1990.

7. Briggs CJ, Capdegelle P, Garner P. Strategies for integrating primary health services in middle- and low-income countries: effects on performance, costs and patient outcomes . The Cochrane Library 2003; (2 (Oxford: Update Software)).

8. Report of the New England Regional Infant Cardiac Program. Pediatrics 1980; 65(2 Pt 2):375-461.

9. Pollack MM, Ruttimann UE, Getson PR. Accurate prediction of the outcome of pediatric intensive care. A new quantitative method. New England Journal of Medicine 1987; 316(3):134-9.

10. Univeristy of York: Centre for health economics. The relationship between hospital volume and quality of health outcomes.

11. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. New England Journal of Medicine 1979; 301(25):1364-9.

12. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann Intern Med 2002; 137(6):511-20.

13. Shackley P, Slack R, Booth A, Michaels J. Is there a positive volume-outcome relationship in peripheral vascular surgery? Results of a systematic review. Eur J Vasc Endovasc Surg 2000; 20(4):326-35.

14. Moher D, Pham B, Klassen TP *et al*. What contributions do languages other than English make on the results of meta-analyses? J of Clin Epi 2000; 53:964-72.

15. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 1998; 52:377-84.

16. University of York NCfRaD. Hospital Volume, Physician Volume, and Patient Outcomes: assessing the evidence. 1995; 2:1-34.

17. Chamberlain JM, Sills MR, Shao C, Joseph JG. The Association of Hospital Type and Volume With Length of Stay and Costs for Appendicitis. Pediatric Research 2000; 47(4. S):1058.

18. Chamberlain J, Teach S, Frey C. The Association of Hospital Volume with Length of Stay and Costs for Appendicitis in California. Emergency Medicine 2001; 469(Poster 30):84A.

19. StillerC, Draper G. Treatment centre size, entry to trials, and survival in acute lymphoblastic leukaemia. Arch Dis Child 1989; 64(5):657-61.

20. Stiller CA. Centralisation of treatment and survival rates for cancer. Archives of Disease in Childhood 1988; 63(1):23-30.

21. Meadows AT, Kramer S, Hopson R, Lustbader E, Jarrett P, Evans AE. Survival in childhood acute lymphocytic leukemia: effect of protocol and place of treatment. Cancer Invest 1983; 1(1):49-55.

22. Albright AL, Sposto R, Holmes E *et al*. Corre lation of neurosurgical subspecialization with outcomes in cildren with malignant brain tumors. Neurosurgery 2000; 47(4):879-87.

23. Danjoux CE, Jenkin DT, McLaughlin J *et al*. Childhood medulloblastoma in Ontario, 1977-1987: population-based results. Medical and Pediatric Oncology 1996; 26:1-9.

24. Duffner PK, Cohen ME, Flannery JT. Referral patterns of childhood brain tumors in the state of Connecticut. Cancer 1982; 50:1636-40.

25. Foreman NK. Variation in survival of children with cancer within a re gion of the United Kingdom. Cancer 1996; 77(4):785-90.

26. Kramer S, Meadows AT, Pastore G, Jarrett P, Bruce D. Influence of place of treatment of diagnosis, treatment, and survival in three pediatric solid tumors. Journal of Clinical Oncology 1984; 2(8):917-23.

27. Sanders BM, Draper GJ, Kingston JE. Retinoblastoma in Great Britain 1969-1980: incidence, treatment, and survival. British Journal of Ophthalmology 1988; 72:576-83.

28. Stiller CA, Eatock EM. Survival from acute non-lymphocytic leukaemia, 1971-1988: a population based study. Archives of Disease in Childhood 1994; 70:219-23.

29. Stiller CA, Eatock EM. Patterns of care and survival for children with acute lympoblastic leukamia diagnosed between 1980 and 1994. Archives of Disease in Childhood 1999; 81:202-8.

30. Chameides L, Galioto F, Fyler DC. The Regional Infant Cardiac Program. A six year evaluation. Connecticut Medicine 1975; 39(11):709-10.

31. Chang RK, Chen AY, Klitzner TS. Factors associated with age at operation for children with congenital heart disease. Pediatrics 2000; 105(5):1073-81.

32. Flyer DC, Parisi L, Berman MA. The regionalization of infant cardiac care in New England. Cardiovasc Clin 1972; 4:339.

33. Hannan EL, Racz M, Kavey RE, Quaegebeur JM, Williams R. Pediatric cardiac surgery: the effect of hospital and surgeon volume on in-hospital mortality. Pediatrics 1998; 101(6):963-9.

34. Jenkins KJ, Newburger JW, Lock JE, Davis RB, Coffman GA, Iezzoni LI. In-hospital mortality for surgical repair of congenital heart defects: preliminary observations of variation by hospital caseload. [see comments.]. Pediatrics 1995; 95(3):323-30.

35. Lundstrom NR , Berggren H, Bjorkhem G, Jogi P, Sunnegardh J. Centralization of pediatric heart surgery in Sweden. Pediatric Cardiology 2000; 21(4):353-7.

36. Norwood W, Dobell A, Freed M, Kirklin JW BE. Intermediate results of the arterial switch repair. J Thorac Cardiovasc Surg 1988; 96:854-63.

37. Scott WA, Fixler DE. Effect of center volume on outcome of ventricular septal defect closure and arterial switch operation. American Journal of Cardiology 2001; 88(11):1259-63.

38. Sollano JA, Gelijns AC, Moskowitz AJ *et al*. Volume-outcome relationships in cardiovascular operations: New York State, 1990-1995. Journal of Thoracic & Cardiovascular Surgery 1999; 117(3):419-28.

39. Spiegelhalter DJ. Mortality and Volume of Cases in Paediatric Cardiac Surgery: Retrospective Study Based on Routinely Collected Data. British Medical Journal 2002; 324(7332):261-4.

40. Stark J, Gallivan S, Lovegrove J *et al*. Mortality rates after surgery for congenital heart defects in children and surgeons' performance. Lancet 2000; 355(9208):1004-7.

41. Linder-Pelz S, Webster MA, Martins J, Greenwell J. Obstetric risks and outcomes: birth centre compared with conventional labour ward. Community Health Studies 1990; 14(1):39-46.

42. Mayfield JA, Rosenblatt RA, Baldwin LM, Chu J, Logerfo JP. The relation of obstetrical volume and nursery level to perinatal mortality. [see comments.]. American Journal of Public Health 1990; 80(7):819-23.

43. Paneth N, Kiely JL, Wallenstein S, Susser M. The choice of place of delivery. Effect of hospital level on mortality in all singleton births in New York City. American Journal of Diseases of Children 1987; 141(1):60-4.

44. Roder D, Chan A, Esterman A. Birthweight-specific trends in perinatal mortality by hospital category in South Australia, 1985-1990. Medical Journal of Australia 1993; 158(10):664-7.

45. Rosenblatt RA, Reinken J, Shoemack P. Is obstetrics safe in small hospitals? Evidence from New Zealand's regionalised perinatal system . Lancet 1985; 2(8452):429-32.

46. Rosenblatt RA, MacFarlane A, Dawson AJ *et al*. The regionalization of perinatal care in Wales and Washington State. American Journal of Public Health 1996; 86(7):1011-5.

47. Verloove-Vanhorick P, Verwey R, Ebeling M, Brand R, Ruys J. Mortality in very preterm and very low birth weight infants according to place of birth and level of care: results of a national collaborative survey of preterm and very low birth weight infants in the Netherlands. Pediatrics 1988; 81(404-411).

48. Sangala V, Dunster G, Bohin S, Osborne JP. Perinatal mortality rates in isolated general practitioner maternity units. [see comments.]. BMJ 1990; 301(6749):418-20.

49. Becerra JE, Perez dS, Smith JC. Evaluation of a regionalized perinatal care system through linked infant birth and death certificates: Puerto Rico, 1980-84. Puerto Rico Health Sciences Journal 1989; 8(2):253-8.

50. Vitsainen K, Gissler M, Hemminki E. Birth outcomes by level of obstetric care in Finland: a catchment area based analysis. Journal of Epidemiology & Community Health 1994; 48(4):400-5.

51. Yeast JD, Poskin M, Stockbauer JW, Shaffer S. Changing patterns in regionalization of perinatal care and the impact on neonatal mortality. American Journal of Obstetrics & Gynecology 1998; 178(1 Pt 1):131-5.

52. LeFevre M, Sanner L, Anderson S, Tsutakawa R. The relationship between neonatal mortality and hospital level. Journal of Family Practice 1992; 35(3):259-64.

53. Clarke M, Mason ES, MacVicar J, Clayton DG. Evaluating perinatal mortality rates: effects of referral and case mix. BMJ 1993; 306(6881):824-7.

54. Cordero L, Zuspan FP. Very low-birth weight infants: five years experience of a regional perinatal program. Ohio Medicine 1988; 84(12):976-8.

55. Gortmaker S, Sobol A, Clark C, Walker DK, Geronimus A. The survival of very low-birth weight infants by level of hospital of birth: a population study of perinatal systems in four states. Am J Obstet Gynecol 1985; 152 (5):517-24.

56. Rosenblatt RA, Mayfield JA, Hart LG, Baldwin LM. Outcomes of regionalized perinatal care in Washington State. Western Journal of Medicine 1988; 149(1):98-102.

57. Black DP, Fyfe IM. The safety of obstetric services in small communities in northern Ontario. CMAJ: Canadian Medical Association Journal 1984; 130(5):571-6.

58. Bowes WAJr. A review of perinatal mortality in Colorado, 1971 to 1978, and its relationship to the regionalization of perinatal services. American Journal of Obstetrics & Gynecology 1981; 141(8):1045-52.

59. Bode MM, O'shea TM, Metzguer KR, Stiles AD. Perinatal regionalization and neonatal mortality in North Carolina, 1968-1994. American Journal of Obstetrics & Gynecology 2001; 184(6):1302-7.

60. Lagoe RJ, Milliren JW, Baader MJ. Impact of selected diagnosis-related groups on regional neonatal care. Pediatrics 1986; 77(5):627-32.

61. Berg CJ, Druschel CM, McCarthy BJ, LaVoie M, Floyd RL. Neonatal mortality in normal birth weight babies: does the level of hospital care make a difference? Am J Obstet Gynecol 1989; 161(1):86-91.

62. Baruffi G, Dellinger WS Jr, Stobino DM, Rudolph A, Timmons RY, Ross A. A study of pregnancy outcomes in a maternity center and a tertiary care hospital. Am J Public Health 1984; 74(9):973-8.

63. Field D, Hodges S, Mason E, Burton P. Survival and place of treatment after premature delivery. [see comments.]. Archives of Disease in Childhood 1991; 66(4 Spec No):408-10.

64. Moster D, Terje LR, Markestad T. Neonatal mortality rates in communities with small maternity units compared with those having larger maternity units. BJOG: an International Journal of Obstetrics & Gynaecology 2001; 108(9):904-9.

65. Moster D, Lie RT, Markestad T. Relation between size of delivery unit and neonatal death in low risk deliveries: population based study. Arch Dis Child Fetal Neonatal Ed 1999;

80(3):F221-5.

66. Siegel E, Gillings D, Campbell S, Guild P. Controlled evaluation of rural regional perinatal care: developmental and neurologic outcomes at 1 year. Pediatrics 1986; 77(2):187-95.

67. International Neonatal Network, Scottish Neonatal Consultants, Nurses Collaborative Study Group. Risk adjusted and population based studies of the outcome for high risk infants in Scotland and Australia. Arch Dis Child Fetal Neonatal Ed 2000; 82(2):F118-23.

68. Hein HA, Burmeister LF. The effect of ten years of regionalized perinatal health care in Iowa, U.S.A. European Journal of Obstetrics, Gynecology, & Reproductive Biology 1986; 21(1):33-48.

69. Hein HA. The quality of perinatal care in small rural hospitals. JAMA 1978; 240(19):2070-2.

70. Fallis G, Dunn E, Hilditch J. Small hospital obstetrics: Is small beautiful? Journal of Rural Health 1988; 4(2):101-17.

71. Campbell DM. Why do physicians in neonatal care units differ in their admission thresholds? Social Science & Medicine 1984; 18(5):365-74.

72. Klein MC, Spence A, Kaczorowski J, Kelly A, Grzybowski S. Does delivery volume of family physicians predict maternal and newborn outcome? CM AJ: Canadian Medical Association Journal 2002; 166(10):1257-63.

73. LeFevre ML. Physician volume and obstetric outcome. Medical Care 1992; 30(9):866-71.

74. Lumley J. The safety of small maternity hospitals in Victoria 1982-84. Community Health Stud 1988; 12(4):386-93.

75. Rozier JCJr. Perinatal mortality in a community hospital after the initiation of a high-risk perinatal program. North Carolina Medical Journal 1980; 41(10):653-6.

76. Walker DJ, Vohr BR, Oh W. Economic analysis of regionalized neonatal care for very low-birth-weight infants in the state of Rhode Island. Pediatrics 1985; 76(1):69-74.

77. Williams RL. Measuring the effectiveness of perinatal medical care. Medical Care 1979; 17(2):95-110.

78. Usher R. Changing mortality rates with perinatal intensive care and regionalization. Seminars in Perinatology 1977; 1(3):309-19.

79. Piekkala PER , Kero Peal. Declining Perinatal Mortality in a Region of Finland, 1968-82. American Journal of Public Health 1985; 75(2):156-60.

80. McCormick MC , Shapiro S, Starfield BH. The Regionalization of Perinatal Services. Summary of the Evaluation of a National Demonstration Program. JAMA 1985; 253(6):799-804.

81. Hein HA, Brown CJ. Neonatal mortality review: a basis for improving care. Pediatrics 1981; 68(4):504-9.

82. Hein HA. Evaluation of a rural perinatal care system. Pediatrics 1980; 66(4):540-6.

83. Hein HA, Lofgren MA. The changing pattern of neonatal mortality in a regionalized system of perinatal care: a current update. Pediatrics 1999; 104(5):1064-9.

84. Peddle LJ, Brown H, Buckley J *et al*. Voluntary regionalization and associated trends in perinatal care: the Nova Scotia Reproductive Care Program. American Journal of Obstetrics & Gynecology 1983; 145(2):170-6.

85. Siegel Earl, Gillings Dennis, Campbell Suzann, Guild Priscilla. A Controlled Evaluation of Rural Regional Perinatal Care: Impact on Mortality and Morbidity. American Journal of Public Health 1985; 75(3):246-53.

86. Shapiro S. Epidemiologic evaluation of large-scale perinatal programs. Mead Johnson Symposium on Perinatal & Developmental Medicine 1982; (20):13-21.

87. Taylor J, Zweig S, Williamson H, Lawhorne L, Wright H. Loss of a rural hospital obstetric unit: a case study. Journal of Rural Health 1989; 5(4):343-52.

88. Callon HF. Regionalizing perinatal care in Wisconsin. Nursing Clinics of North America 1975; 10(2):263-74.

89. Hoekstra R, Fangman J, Perkett E, Brasel D, Knox GE. Regionalization of perinatal care: results of a cooperative community based program. Minnesota Medicine 1981; 64(10):637-40.

90. Bakketeig LS , Hoffman HJ, Sternthal PM. Obstetric service and perinatal mortality in Norway. Acta Obstet Gynecol Scand Suppl 1978; 77:3-19.

91. Field D, Draper ES. Survival and place of delivery following preterm birth: 1994-1996. Archives of Disease on Childhood - Fetal Neonatal Edition 1999; 80(2):F111-F114.

92. Goodman DC, Fisher ES, Little GA, Stukel TA, Chang CH, Schoendorf KS. The relation between the availability of neonatal intensive care and neonatal mortality [see comments]. New England Journal of Medicine 2002; 346(20):1538-44.

93. Horbar JD, McAuliffe TL, Adler SM *et al*. Variability in 28-day outcomes for very low birth weight infants: an analysis of 11 neonatal intensive care units. Pediatrics 1988; 82(4):554-9.

94. Horbar JD, Badger GJ, Lewit EM, Rogowski J, Shiono PH. Hospital and patient characteristics associated with variation in 28-day mortality rates for very low birth weight infants. Vermont Oxford Network. [see comments.]. Pediatrics 1997; 99(2):149-56.

95. Paneth N, Kiely JL, Wallenstein S, Marcus M, Pakter J, Susser M. Newborn intensive care and neonatal mortality in low-birth-weight infants: a population study. N Engl J Med 1982; 307(3):149-55.

96. Phibbs CS, Bronstein JM, Buxton E, Phibbs RH. The effects of patient volume and level of care at the hospital of birth on neonatal mortality. JAMA 1996; 276(13):1054-9.

97. The UK Neonatal Staffing Study Group. Patient volume, staffing, and workload in relation to risk-adjusted outcomes in a random stratified sample of UK neonatal intensive care units: a prospective evaluation. [see comments.]. Lancet 2002; 359(9301):99-107.

98. DeNicola LK. Multi-institutional approach to regionalization of pediatric critical care. Seven year experience. Journal of the Florida Medical Association 1992; 79(3):159-64.

99. Lopez-Herce J, Sancho L, Martinon JM. Study of paediatric intensive care units in Spain. Spanish Society of Paediatric Intensive Care. [see comments.]. Intensive Care Medicine 2000; 26(1):62-8.

100. Pearson G, Shann F, Barry P *et al*. Should paediatric intensive care be centralised? Trent versus Victoria. Lancet 1997; 349(9060):1213-7.

101. Pearson G, Barry P, Timmins C, Stickley J, Hocking M. Changes in the profile of paediatric intensive care associated with centralisation. Intensive Care Medicine 2001; 27(10):1670-3.

102. Pollack MM, Cuerdon TT, Patel KM *et al*. Impact of quality-of-care factors on pediatric intensive care unit mortality. [see comments.]. JAMA 1994; 272(12):941-6.

103. Pollack MM, Alexander SR, Clarke N, Ruttimann UE, Tesselaar HM, Bachulis AC. Improved outcomes from tertiary center pediatric intensive care: a statewide comparison of tertiary and nontertiary care facilities. [see comments.]. Critical Care Medicine 1991; 19(2):150-9.

104. Tilford JM, Simpson PM, Green JW, Lensing S, Fiser DH. Volume-outcome relationships in pediatric intensive care units. Pediatrics 2000; 106(2 Pt 1):289-94.

105. Cox K, Rodriguez-Baez N, Nasr A, Esquivel C. Mortality rate correlated with the number of pediatric liver transplants performed at a center. Transplantation Proceedings 2001; 33(1-2):1512-3.

106. Hosenpud JD , Breen TJ, Edwards EB, Daily OP, Hunsicker LG. The effect of transplant center volume on cardiac transplant outcome. A report of the United Network for Organ Sharing Scientific Registry. JAMA 1994; 271(23):1844-9.

107. Schurman SJ , Stablein DM, Perlman SA, Warady BA. Center volume effects in pediatric renal transplantation. A report of the North American Pediatric Renal Transplant Cooperative Study. Pediatric Nephrology 1999; 13(5):373-8.

108. Cooper A, Barlow B, DiScala C, String D, Ray K, Mottley L. Efficacy of pediatric trauma care: results of a population-based study. Journal of Pediatric Surgery 1993; 28(3):299-303.

109. Cosentino CM, Barthel MJ, Reynolds M. The impact of level 1 pediatric trauma center designation on demographics and financial reimbursement. Journal of Pediatric Surgery 1991; 26(3):306-9.

110. Konvolinka CW, Copes WS, Sacco WJ. Institution and per-surgeon volume versus survival outcome in Pennsylvania's trauma centers. American Journal of Surgery 1995; 170(4):333-40.

111. Nakayama DK , Copes WS, Sacco W. Differences in trauma care among pediatric and nonpediatric trauma centers. Journal of Pediatric Surgery 1992; 27(4):427-31.

112. Potoka DA, Schall LC, Gardner MJ, Stafford PW, Peitzman AB, Ford HR. Impact of pediatric trauma centers on mortality in a statewide system. Journal of Trauma -Injury Infection & Critical Care 2000; 49(2):237-45.

113. Tepas JJI, Patel JC, DiScala C, Wears RL, Veldenz HC. Relationship of trauma patient volume to outcome experience: can a relationship be defined? Journal of Trauma -Injury Infection & Critical Care 1998; 44(5):827-30.

114. VanRooyen MJ, Sloan EP, Barrett JA, Smith RF, Reyes HM. Outcome in an urban pediatric trauma system with unified prehospital emergency medical services care. Prehospital & Disaster Medicine 1995; 10(1):19-23.

115. Witt PD, Wahlen JC, Marsh JL, Grames LM, Pilgram TK. The effect of surgeon experience on velopharyngeal functional outcome following palatoplasty: is there a learning curve? Plastic

& Reconstructive Surgery 1998; 102 (5):1375-84.

116. Albright AL , Pollack IF, Adelson PD, Solot JL. Outcome data and analysis in pediatric neurosurgery. Neurosurgery 1999; 45(1):101-6.

117. Baumer JH, Hunt LP, Shield JP. Audit of diabetes care by caseload. Archives of Disease in Childhood 1997; 77(2):102-7.

118. Bratton SL, Haberkern CM, Waldhausen JH, Sawin RS, Allison JW. Intussusception: hospital size and risk of surgery. Pediatrics 2001; 107(2):299-303.

119. Snow BW, Catwright PC, Young MD. Does surgical subspecialization in pediatrics provide high-quality, cost-effective patient care? Pediatrics 1996; 97(1):14-7.

120. Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. JAMA 1998; 280(20):1747-51.

121. Kelly JV, Hellinger FJ. Heart disease and hospital deaths: an empirical study. Health Services Research 1987; 22:369-95.

122. Boulware LE , Daumit GL, Frick KD, Minkovitz CS. Quality of clinical reports on behavioral interventions for hypertension. Prev Med 2002; 34(4):463-75.

123. Hennekens C , Buring J. Epidemiology in Medicine. Lippincott Williams & Wilkins.

124. Szczech LA, Coladonato JA, Owen WF. Study designs and their potential influence on conclusions. Semin Dial 2002; 15(3):207-11.

125. Hannan EL, Kilburn H Jr, Racz M, Shields E, Chassin MR. Improving the outcomes of coronary artery bypass surgery in New Yrok States. JAMA 1994; 271:761-6.

126. Hughes RG, Hunt SS, Luft HS. Effects of surgeon volume and hospital volume on quality of care in hospitals. Medical Care 1987; 25:489-503.

127. Deutchman ME, Sills D, Connor PD. Perinatal outcomes: a comparison between family physicians and obstetricians. J Am Board Fam Pract 1995; 8(6):440-7.

128. Smyth RL. Research with children. Paediatric practice needs better evidence--gained in collaboration with parents and children. BMJ 2001; 322(7299):1377-8.

129. Chang RK, Klitzner TS. Can regionalization decrease the number of deaths for children who undergo cardiac surgery? A theoretical analysis. Pediatrics 2002; 109(2):173-81.

Appendix I: Tables

Appendicitis

Author / Year / Country	Objective	Study Characteristics	Case Mix	Results	Author's Conclusions
Chamberlain JM 2000 United States	To test the hypotheses that hospital type, annual pediatric volume, and the presence of residency training programs are associated with lower costs and shorter length of hospital stay for appendicitis.	Data collection: retrospective Design: observational Primary outcome: cost and length of stay Study period: 1995 Comparisons: high volume (discharge >2% of state total) vs low volume (discharge =2% of state total) Study population: New York State Data source: New York Statewide Planning and Resource Cooperative System Age groups: <18 years	Yes Demographic: age, sex, race, Medicaide insurance, residency in New York City Stage/severity of illness: complicated appendicitis (peritonitis or abscess)	NYC Complicated Cases Costs: High volume= $3889 Low volume= $4927 (p=0.0003) Length of stay: High volume= 4.9 days Low volume= 6.3 days (p=0.0004) NYC Simple Cases High volume= $2143 Low volume= $2612 (p=0.0001) Length of stay: High volume= 2.7 days Low volume= 2.9 days (p=0.002) Only significant difference for non-NYC cases was for length of stay of complicated cases: High volume= 4.3 days Low volume= 5.5 days (p=0.0001)	Decreased length of stay and costs for pediatric appendicitis are significantly associated with high volume of cases. Case-mix: Level II

Chamberlain JM 2001 United States	To test the hypothesis that higher annual pediatric volume is associated with similar findings of lower costs and shorter length of hospital stay for appendicitis in California.	Data collection: retrospective Design: observational Primary outcome: survival rate Study period: 1995 Comparisons: high volume (discharge >10 patients a day) vs low volume (discharge =10 patients a day) Study population: California Data source: California statewide discharge database Age groups: <18 years	Yes Demographic: age, race, insurance Stage/severity of illness: complicated appendicitis (death, peritonitis, or abscess)	Complicated Cases Mean Adjusted Costs High volume: $15 223 Low volume: $12 360 (p=0.0001) Mean Adjusted LOS High volume: 5 days Low volume: 4 days (p=0.0001) Simple cases Mean Adjusted Costs High volume: $7798 Low volume: $6768 (p=0.0001) Mean Adjusted LOS High volume: 2 days Low volume: 2 days (NS)	California high volume hospitals treated complicated pediatric appendicitis patients less efficiently (higher costs and longer length of stay) than low volume hospitals. Case-mix: Level II

Cancer

Author / Year / Country	Objective	Study Characteristics	Case Mix	Results	Author's Conclusions
Albright AL 2000 United States	To evaluate the association between the type of neurosurgeon (general or pediatric) and the extent of tumor removal or the frequency of complications in children undergoing malignant brain tumor resections.	Data collection: retrospective Design: observational Primary outcome: extent of tumor removal or complications Study period: 1985-1992 Comparisons: General (6 mo. of pediatric training in residency and occasionally perform pediatric operations) vs Pediatric (member of American Society of Pediatric Neurosurgeons or designated pediatric neurosurgeon in their centre but not a member of ASPN) Study population: patients enrolled in three recent Children's Caner Group studies Data source: neuorsurgical data sheets and operative reports Age groups: not stated	No	Extent of Tumor Removal of >90% ASPN members: 33% Designated pediatric NS: 31% General NS: 42% Neurological Complications ASPN members: 17.7% Designated pediatric NS: 32.0% General NS: 22.5%	Pediatric neurosurgeons are more likely to extensively remove malignant pediatric brain tumors than general surgeons and ideally children with brain tumors should be operated on by neurosurgeons who perform numerous brain tumor operations per year. Case-mix: Level 0

Danjoux CE 1986 Canada	To review all 108 children not more than 16 years old, diagnosed with posterior fossa medulloblastoma in Ontario to provided outcome data and compare the rates of relapse-free survival and examine whether these differences in outcome can be explained by other prognostic factors.	Data collection: retrospective Design: observational Primary outcome: relapse-free survival Study period: 1977-1987 Comparisons: large vs medium vs small centre Study population: Ontario Cancer Registry Data source: discharge summaries, pathology reports, specialized cancer centres, and death certificates Age groups: <16 years	No	Five-Year Relapse Free Survival Large centre: 65% Medium centre: 28% Small centre: 8%	For relatively rare diseases, such as medulloblastoma, which require complex treatment; a minimal number of cases should be seen to ensure optimal expertise. Case-mix: Level 0
Duffner PK 1982 United States	To determine the referral patterns of children with brain tumors in Connecticut and to determine if the survival rates of children treated at university cancer centres differed from those treated at community hospitals.	Data collection: retrospective Design: observational Primary outcome: survival Study period: 1968-1979 Comparisons: university cancer centres vs community hospitals Study population: Connecticut Tumor Registry Data source: patient and computer records Age groups: 0-18 years	No	Projected Survival at 5-Years Medulloblastomas University centre: 74% Community centre: 29% ($p<0.01$) Brainstem Gliomas University centre: 40% Community centre: 0% ($p<0.01$) Survival was similar for cerebellar astrocytomas, midline astroyctomas, astrocytomas (Grade I, II, III, and IV), ependymomas, glioblastomas.	Medulloblastomas and brainstem glimas had significantly higher projected survivals for children treated in university centres compared to those receiving all t heir treatment at community hospitals. For the other types of brain tumors, survival was similar. Case-mix: Level 0

Foreman NK 1996 United Kingdom	To look at the variation in survival in a large region of the United Kingdom and some possible reasons for it.	Data collection: retrospective Design: observational Primary outcome: survival Study period: 1976-1985 Comparisons: large (>6 patients a year) vs small (<6 patients a year) Study population: five counties in southwest UK Data source: hospital inpatient surveys, ward admission books, correspondence with subspecialty units, note collection of various departments, Regional Pediatric Oncology Unit, tumor registries, and pathology department records Age groups: <15 years	Yes Stage/severity of illness: type of malignancy	Five-Year Survival (1976-1985) CNS tumor Large centre: 58.4% Small centre: 41.4% (p=0.03) Leukemia Large centre: 53.4% Small centre: 56.5% (p=0.90) Other cancers Large centre: 66.1% Small centre: 60.1% (p=0.24)	Survival in children with CNS tumors was significantly better in large centres for the first half of the study period (1976-1980). There is no significant difference in survival based on hospital size for patients with leukemia or other cancers. Case-mix: Level II

Kramer S 1984 United States	To examine differences between cancer centres and non-cancer centres in terms of management procedures and outcomes for three pediatric solid tumors: Wilms' tumor, rhabdomyosarcoma, and medulloblastoma.	Data collection: retrospective Design: observational Primary outcome: three-year disease-free survival Study period: 1970-1979 Comparisons: Cancer centre (hospital with a full-time pediatric hematologist/oncologist) vs non-cancer centre (not defined) Study population: Greater Delaware Valley Data source: Greater Delaware Valley Pediatric Tumor Registry Age groups: 0-14 years	No	Disease-free Survival Rhabdomyosarcoma (1970-1979) Cancer centre: 48% Non-cancer centre: 10% (p=0.002) Wilms' tumor (1975-1979) Cancer centre: 85% Non-cancer centre: 71% (NS) Operative Death Medulloblastoma (1970-1974) Cancer centre: 0 Non-cancer centre: 8 (p=0.02) There was no difference between 1975-1979.	Three-year disease-free survival was significantly better for rhabdomyosarcoma treated in cancer centres; survival was no different for Wilms' tumor. There was an increase in operative deaths in non-cancer centres for medulloblastoma in 1970-1974. Case-mix: Level 0
Meadows AT 1983 United States	To determine the effect of place and type of initial treatment on survival for acute lymphocytic leukemia (ALL).	Data collection: retrospective Design: observational Primary outcome: survival Study period: 1970-1975 Comparisons: cancer centre vs non-cancer centre Study population: 31-counties surrounding Philadelphia Data source: Greater Delaware Valley Pediatric Tumor Registry Age groups: <15 years	No	Survival Cancer centre: 58% Non-cancer centre: 40%	There was a significant improved survival rate among patients treated in a cancer centre. Case-mix: Level 0

Sanders BM 1988 United Kingdom	Patients with retinoblastoma diagnosed from 1969-1980 have been followed up for periods of up to 17 years.	Data collection: retrospective Design: observational Primary outcome: survival Study period: 1969-1980 Comparisons: London hospitals working jointly as the major UK referral centre vs other specialist eye hospitals vs remaining hospitals Study population: England, Scotland, and Wales Data source: National Registry of Childhood Tumors Age groups: up to 17 years	Yes Demographic: age Stage/severity illness: stage of cancer	After controlling for stage of cancer and age, survival rate in unilateral cases was better at major centres, followed by specialist eye hospitals ($p=0.014$). No difference existed for bilateral cases.	Children re ferred to units specializing in the treatment of retinoblastoma have a higher three-year survival rate than those treated at other hospitals. Case-mix: Level II
Stiller CA 1988 United Kingdom	To describe the increased centralization of the care of children with cancer since 1977, and compare survival rates for children with some of the principal types of childhood neoplasm who were treated at pediatric oncology centres with those of children treated elsewhere.	Data collection: retrospective Design: observational Primary outcome: survival rates Study period: 1977-1984 Comparisons: pediatric oncology hospital vs other teaching hospital vs other non-teaching hospital Study population: UK Children's Cancer Study Group Data source: UK Children's Cancer Study Group national registry of childhood malignant diseases for Great Britain Age groups: <15 years	Yes Demographic: age, year of diagnosis	Of the eight diagnostic groups, four (Acute non-lymphoblastic leukemia, Non-Hodgkin's lymphoma, Ewing's tumor, and Rhabdomyosarcoma) had significantly better three-year survival rates when treated at a pediatric oncology centre.	The results underline the importance of referring patients to specialist centres where they may benefit as early as possible for the latest advances in treatment. Case-mix: Level I

Stiller CA 1989 United Kingdom	To examine trends in referral rates to specialist centres and in the proportions of children included in trials, and the survival rates for children treated both inside and outside the trials and at centres of different sizes.	Data collection: retrospective Design: observational Primary outcome: survival rates Study period: 1971-1982 Comparisons: high volume (=6) vs medium (1-5) vs low (<1) Study population: UK centres treating patients with acute lymphoblastic leukemia Data source: National Registry of Childhood Tumors Age groups: <15 years	Yes Demographic: age	When compared to low volume, high volume centres had better survival rates. This trend is consistent across time, however there is less disparity. In 1971-73, the high volume survival rate was 46% and it was 22% in low volume. Between 1980-1982, survival rate in high volume was 67% and 58% in low volume hospitals.	The survival rate was highest children treated in high volume centres. Case-mix: Level I
Stiller CA 1994 United Kingdom	To investigate possible prognostic factors, including those related to patterns of medical care, for acute non-lympothmatic leukemia in childhood in a large population based series and to document survival rates since 1985.	Data collection: retrospective Design: observational Primary outcome: survival Study period: 1971-1988 Comparisons: teaching hospital vs other hospitals Study population: National Registry of Childhood Tumors Data source: National Registry of Childhood Tumors Age groups: <15 years	Yes Demographic: age	Survival Rate Teaching hospitals versus other hospitals: $?^2 = 5.71$, $p<0.05$ When stratified by trial and non-trial children, there is no survival difference amongst treatment centres.	Children treated at teaching hospitals had a higher survival rate than those treated elsewhere, allowing for age. When trial and non-trial children were analyzed separately, there was no significant difference in survival between types of treatment centre for either group. Case-mix: Level I

Stiller CA 1999 United Kingdom	To document survival rates after acute lymphoblastic leukemia (ALL) to assess effects of prognostic factors at presentation, and to investigate the relation of survival to patterns of care organization.	Data collection: retrospective Design: observational Primary outcome: actuarial survival rates Study period: 1980-1994 Comparisons: Mean annual volume of <1, 1 -5, or >6 patients a year or UKCCSG (treated at pediatric oncology centre) vs non-UKCCSG hospitals Study population: National Registry of Childhood Tumors Data source: National Registry of Childhood Tumors Age groups: <15 years	No	5-year % Survival 1980-1984 <1/year: 59 1-5: 66 >6: 69 1985-1989 <1/year: 76 1-5: 79 >6: 73 1990-1994 <1/year: 78 1-5: 80 >6: 82 *differences in survival are not significant	Survival did not vary significantly between hospitals with different numbers of patients during any part of the study period. Case-mix: Level 0

Cardiac

Author / Year / Country	Objective	Study Characteristics	Case Mix	Results	Author's Conclusions
Chameides L 1975 United States	To give a general review of the data accumulated in the past 6 years of the New England Regional Infant Cardiac Program	Data collection: retrospective Design: observational Primary outcome: survival rates Study period: 1968-1973 Comparisons: large centre A (14 patients) vs large centre B (31) vs large centre C (26) vs combined small centres (40 in 7 centres) Study population: New England (Connecticut, Maine, Rhode Island, Massachusetts, New Hampshire, and Vermont) Data source: New England Regional Infant Cardiac Program database Age groups: <1 year	No	Mortality Large centre A: 57% Large centre B: 29% Large centre C: 54% Combined small centres: 30%	There appears to be some centre-to-centre differences in mortality, but this is not related to centre size. Case-mix: Level 0

Chang RR 2000 United States	To identify factors associated with, or influencing the age at which children undergo definitive surgical repair of the most common forms of coronary heart disease.	Data collection: retrospective Design: observational Primary outcome: age of definitive surgical repair Study period: 1995-1996 Comparisons: n/a Study population: California children Data source: discharge data from the California Office of Statewide Health Planning and Development (OSHPD) database Age groups: <18 years	Yes Stage/severity of illness: type of CHD	Median Age and Case Volume Correlation ASD: r= -0.37 VSD: r= -0.49 TOF: r= -0.63 AVC: r= -0.17 All p-values <0.01	Centres with higher surgical case volume were more likely to operate at a younger age. Negative correlation between median age and case volume of surgical centre was found for repair of ASD, VSD, TOF, and AVC. Case Mix: Level II
Flyer DC 1972 United States	To report activities and results of the New England Regional Infant Cardiac Program.	Data collection: retrospective and prospective Design: observational Primary outcome: not stated Study period: 1968 (baseline), 1969 (1st program year), 1970 (2nd program year) Comparisons: 11 centres divided into 8 "hospitals with smaller experience" and 3 large hospitals Study population: New England Data source: database developed for the program and medical chart s Age groups: <1 year	No	Mortality (excluding associated anomalies and low birth weight) 3 large hospitals: 38% 8 small hospitals: 30%	The first program year yielded 50 % increase in the number of babies (better identification as a result of the program). The age of referral to the pediatric cardiologist is becoming progressively younger. The age distribution at surgery indicates that over 50% of cardiac surgery was accomplished at less than 2 months of age in the three hospitals with the largest experience, while in the eight hospitals with more limited patient material 37% of operations were performed on infants less than 2 months old. Mortality for cardiac surgery was 38% in 3 large hospitals and 30% in 8 smaller hospitals (overall 36%) - authors claim not significant. Case-mix: Level 0

Hannan EL 1998 United States	To examine the relationship between annual provider (hospital and surgeon) volume of pediatric cardiac surgery and in-hospital mortality.	Data collection: retrospective Design: observational Primary outcome: mortality rates Study period: 1992-1995 Comparisons: hospital: =100 cardiac surgeries per year vs <100 physician: =75 procedures/year vs <75 Study population: New York State Data source: New York's Cardiac Surgery Reporting System Age groups: not stated	Yes Demographic: age Stage/severity of illness: four procedure complexity categories Co-morbidities: several clinical risk factors	Risk-adjusted Mortality Rates =100 hospital and =75 physician: 5.45 <100 hospital and =75 physician: 7.45 =100 hospital and <75 physician: 8.47 <100 hospital and <75 physician: 8.94	Annual hospital volume and annual surgeon volume were both significantly related to inpatient mortality rates, even after controlling for patient age and several clinical risk factors in addition to procedure complexity. Case-mix: Level III
Jenkins KJ 1995 United States	To examine the impact of hospital caseload on in-hospital mortality for pediatric congenital heart surgery.	Data collection: retrospective Design: observational Primary outcome: in-hospital deaths Study period: 1991-1995 Comparisons: >300 case/yr vs 101-300 vs 10-100 vs <10 Study population: California and Massachusetts Data source: computerized statewide hospital discharge abstract databases Age groups: <18 years	Yes Demographic: sociodemographic variables Stage/severity of illness: four procedure complexity categories	Adjusted OR for In-hospital Death >300: reference 101-300: 3.0 (1.8, 4.9) 10-100: 2.9 (1.6, 5.3) <10: 7.7 (1.6, 37.8)	For children with a congenital heart defect who underwent surgery in California in 1988 or Massachusetts in 1989, the risk of dying in-hospital was much lower if the surgery was performed at an institution performing >300 cases annually. Case-mix: Level II

Lundstrom NR 2000 Sweden	To report the effects of centralization of pediatric cardiac care in Sweden, especially in relation to mortality after surgery.	Data collection: retrospective Design: observational Primary outcome: 30-day mortality Study period: 1988-1991 (pre-centralization) and 1995-1997 (post-centralization) Comparisons: 2 centres after centralization vs 4 centres prior to centralization Study population: Sweden Data source: National Board of Health Age groups: not stated	Yes Stage/severity of illness: open heart surgery was graded according to severity	30-day Mortality Open-heart surgery Post-centralization: 1.9% Pre-centralization: 9.5% Closed-heart surgery: Post-centralization: 2.7% Pre-centralization: 2.0%	Heart surgery was concentrated to the two centres with the lowest surgical mortality, and the reduction in surgical mortality was observed over a short period of time, making it likely that the centralization of the surgical activity promoted the improved results. Case-mix: Level II

New England Regional Infant Cardiac Program- NERICP 1980 United States	The details of specific achievements attributable to NERICP are presented. The fundamental purpose is to document the experience through the first birthday of a large group of critically ill infants.	Data collection: retrospective pre-program, prospective post-program Design: observational Primary outcome: 30-day mortality surgical mortality and mortality at first birthday Study period: 1969-1974 (pre-centralization) and 1995-1997 (post-centralization) Comparisons: five hospitals offering a full cardiac service and three hospitals offering limited surgical services (1974) vs 11 participating hospitals in 1969 Study population: members of NERICP Data source: state vital statistics Age groups: <1 year	Yes Stage/severity of illness: open vs closed surgery	Crude 30-day Mortality for Open/Closed Heart Surgery: 1969: 50 / 32 1970: 57 / 28 1971: 60 / 32 1972: 52 / 29 1973: 50 / 36 1974: 39 / 29	In spite of differences in numbers of new patients and variations in surgical practice, the individual hospitals managed infants with heart disease with relatively similar results; authors believe this was a major contribution of NERICP, which resulted in similar survival no matter where the baby was born. Case-mix: Level II

Norwood WI 1988 United States	To define more fully some of the early results of the arterial switch repair.	Data collection: prospective Design: observational Primary outcome: mortality and survival (hospital and post-hospital deaths) Study period: Jan 1, 1985-June 1, 1987 Comparisons: =50 repairs per year vs 40-49 vs 30-39 vs 20-29 vs 10-19 vs <10 Study population: 20-institution study conducted in North America Data source: Data Collection and Analysis Centre in Birmingham, Alabama Age groups:<15 days	No	Death =50: 9% 40-49: no institutions in category 30-39: no institutions in category 20-29: 9% 10-19: 36% <10: 55%	Mortality rate varied inversely with the number of arterial switch repairs performed. Case-mix: Level 0
Scott WA 2001 United States	To determine the strength of association between surgical volume and length of stay.	Data collection: retrospective Design: observational Primary outcome: length of stay Study period: 1992-1996 Comparisons: VSD: 1679 patients (median 62/centre; range 17-168) and ASO: 494 patients (median 14/centre; range 4-51) Study population: 24 centres that contributed to the Pediatric Cardiac Care Consortium Registry Data source: Pediatric Cardiac Care Consortium	Yes Demographic: age, weight Co-morbidity: trisomy 21, co-existing heart defects Stage/severity of illness: duration of preoperative hospital stay	Association between centre volume and LOS in multivariate modeling: VSD (p=0.738) ASO (p=0.010)	For complicated procedures, length of stay tends to be shorter in centres with higher volume. This relationship was not observed for uncomplicated procedures. Case-mix= Level III

		Registry Age groups: 1 week- 18 years for VSD; not stated for ASO			
Sollano JA 1999 United States	To examine the relationship between hospital volume and in-hospital deaths in coronary artery bypass grafting, elective repair of abdominal aortic aneurysms, and repair of congenital cardiac defects.	Data collection: retrospective Design: observational Primary outcome: in-hospital deaths Study period: 1990-1995 Comparisons: hospital volume/100 cases Study population: New York State (adults and children; data for children presented separately) Data source: Statewide Planning and Research Cooperative System; for pediatrics, clinical data are routinely collected by the Cardiac Advisory Committee Age groups: <18 years	Yes Demographic: age Stage/severity of illness: surgical complexity categories developed by Jenkins	Case Severity Adjusted OR Hospital volume/100= 0.944 (0.921,0.967)	Pediatric patients undergoing cardiac surgical procedures in general demonstrate a significant inverse relationship between hospital volume and hospital death. This relationship is most pronounced in neonates for whom hospitals with volumes below 50 patients per year had mortality rates that were 2 times higher than in the highest volume hospitals. Case-mix: Level II

Spiegelhalter DJ 2001 United Kingdom	To determine whether mortality between 1991 and 1995 in hospitals in England carrying out surgery for congenital heart disease was associated with the annual volume of cases and to estimate the extent to which an association could explain the apparent divergent mortality at Bristol Royal Infirmary.	Data collection: retrospective Design: observational Primary outcome: 30 day mortality Study period: 1991-1995 Comparisons: the relative change in odds of death per additional patient per year Study population: UK Data source: hospital episode statistics and cardiac surgical register Age groups: not stated	Yes Demographic: age Stage/severity of illness: case-mix thru procedure group	Change in Odds of Death (%): All open operations Aged <1: -0.38 (-0.71,-0.004) Aged =1: -0.20 (-0.81,0.46) Open operations (stratified by type of procedure) Aged <1: -0.56 (-0.92,-0.20) Aged =1: -0.22 (-0.60,0.18) Closed operations Aged <1: -0.28 (-1.60, 1.24) Aged =1: -9.01 (-9.82, 0.004) These results exclude Bristol data; including data from Bristol increases the association and its significance. The results presented here are based on hospital episode statistics; results were consistent when using cardiac surgical register	Mortality in children aged less than 1 year old who underwent open heart surgery is significantly related to the volume of cases, even when data from Bristol are excluded. Case-mix: Level II
Stark J 2000 United Kingdom	To gather and report about surgeons' performance for one year to provide information for clinicians, researchers, policy makers, and the general public.	Data collection: retrospective Design: observational Primary outcome: in-hospital death Study period: April 1, 1997 to March 31, 1998 Comparisons: 11 physicians with a volume range of 19-235 Study population: 5 "selected hospitals" in the UK Data source: computerized database from each of the selected centers Age groups: >1 year and no upper age limit	Yes Stage/severity of illness: type of procedure	Although some mortality rates are numerically large, no reliable conclusions can be drawn.	The participating departments seemed to reach high standards of care for children with congenital heart defects, although more data would be needed to assess performance of individual surgeons. Case-mix: Level II

Neonatal: Level of Care

Author / Year / Country /	Objective	Study Characteristics	Case Mix	Results	Author's Conclusions
Baruffi G 1984 US	To investigate pregnancy outcomes between two groups of pregnant women who were similar in sociodemographic characteristics but gave birth in different institutions, and to determine if any differences noted between these groups were due to differences in medical-obstetric risk of women delivering at the two institutions.	Data collection: retrospective Design: observational Primary outcome: neonatal morbidity and mortality (in hospital death) Study period: 1977-1978 Comparisons: primary care, family orientated hospital for low risk births vs tertiary care with NICU Study population: Philadelphia Data source: not stated	Yes Demographic: race, maternal age, parity, prior pregnancy loss, maternal education Stage/severity of illness: Medical-obstetrical risk (based on prenatal and intrapartum risk factors), method of delivery	Difference in Neonatal Mortality Rate, by Perinatal (PN) and Intrapartum (IP) Risk: Low PN/low IP: 3.0 (-2.8,8.7) High PH/low IP:0 Low PN/high IP: -4.8 (-22.9,13.3) High PN/High IP: -30.1 (-56.1,-4.1) (rates were 0 at primary centre, 30.1 at tertiary centre)	The percentage of infants with one-minute Apgar scores <7 or requiring resuscitation at birth was greater at the primary care centre, but the percentage of infants with 5-minute Apgar scores <7 and neonatal mortality did not differ between the two institutions. The study results suggest that care delivered at a primary care centre is safe with regard to the evaluative criteria used. Case mix: Level II

Becerra JE 1989 Puerto Rico	By using linked infant birth-death certificates from 1980 through 1984 to evaluate the regionalization of perinatal health services operating in Puerto Rico.	Data collection: retrospective Design: observational Primary outcome: neonatal death (<28 days) Study period: 1980-1984 Comparisons: primary centres vs regional hospitals vs medical hospitals Study population: Puerto Rico Data source: Puerto Rico Department of Health provided linked birth-death certificates	Yes Demographic: maternal age education, marital status Stage/severity of illness: birth weight, 5 minute Apgar score, congenital anomalies Co-morbidities: pregnancy complications	Authors found more neonatal deaths than expected in level A and B hospitals after adjusting for important confounders (numbers not presented in report). When limiting analysis of level B hospitals to those with NICUs, the excess of observed to expected deaths was larger and statistically significant (numbers not reported).	The regionalization plan in Puerto Rico requires reevaluation; recommendations are put forward by the authors. Case-mix: Level III
Berg CJ 1989 United States	To examine the impact that the level of neonatal care available at the hospital of delivery had on neonatal mortality in infants with birth weights of 2500 grams and greater.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality Study period: 1969-1982 Comparisons: Level III (NICU) vs Level II (intermediate level of care) vs Level I (ability to stabilize patients for transport) Study population: Georgia Data source: Office of Vital Records	Yes Demographic: race, maternal age Stage/severity of illness: prior fetal loss, prior low birth weight delivery, complications unrelated to pregnancy, multiple gestation (variables were weighted and summed to yield a risk score)	RR of Neonatal Death at Level I vs Level III hospital by Risk Status: Low: reference Moderate: 1.2 (95%CI not reported) High: 1.6 (1.1, 2.4)	Authors suggest prenatal risk assessment for all women and referral of high-risk women to level II or III hospital for delivery even at term. Case-mix: Level II

Black DP 1984 Canada	To determine the relationship between the level of services provided in small communities and the obstetric and neonatal outcomes.	Data collection: retrospective Design: observational Primary outcome: perinatal loss rate (number of stillbirths and deaths within 28 days of birth per 1000 newborn admissions) Study period: Apr 1980 - Mar 1982 Comparisons: Level 0 (no elective deliveries) vs Level IA (elective deliveries, but anesthesia not available) vs Level IB (anesthesia available, but =5 cesarean sections per year) vs Level IC (cesarean sections done regularly but specialists in obstetrics or pediatrics not on staff) vs Level ID (at least 1 specialist in obstetrics or pediatrics on staff, fewer than 1000 deliveries per year) vs Level II (=2 specialists in both obstetrics and pediatrics on staff, >1000 deliveries per year) Study population: Northern Ontario Data source: Hospital Medical Records Institute	No	Perinatal Deaths (per 1000 newborn admissions) Level 0: 13.89 (2.87, 40.60) Level IA: 13.73 (7.85, 22.30) Level IB: 14.66 (9.58, 20.80) Level IC: 10.43 (7.61, 13.96) Level ID: 12.72 (8.06, 19.09) Level II: 12.13 (10.37,13.89)	Relatively safe perinatal care is being provided in the small hospitals of northern Ontario. The judicious transfer of patients to larger centres, results in care that is as safe as that provided to patients who have access to the larger, specialized hospitals. While the results do not support greater centralization of obstetric services, neither do they mean that a regional program should not be developed. Such a program could bring beneficial changes through better communication, education and transportation. Case-mix: Level 0

Bode MM 2001 United States	To analyze trends across time in the regionalization of low-birth-weight births and time trends for the association between regionalization and decreased neonatal mortality.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality (death within the first 28 days of life) Study period: 1968-1994 Comparisons: Level III (routinely provide care for mothers and neonates of all risk categories) vs Level II (care for prematurely born neonates and infants born before 34 weeks' gestation whose acute neonatal problems had resolved) vs Level I (obstetric and neonatal care is offered) Study population: North Carolina Data source: Birth and Consolidated Birth/Infant Death files provided by the North Carolina State Centre for Health Statistics	Yes Demographic: maternal age, years of education, marital status, infant race, gestational age, epoch of birth Stage/severity of illness: birth weight	For 1968-1974, neonates >1501 g were at lower risk of dying if born in Level III (OR 0.80 Level I, 0.96 Level II); neonates = 1500 g didnot show same benefit of birth in level III centre (OR ranged from 0.84 to 2.71). After 1974, neonates =1500 g were consistently at lower risk of dying if born in level III centre (OR ranged from 0.31 to 0.92). There was no difference between level II and level I for neonates =1500 g. The survival advantage of birth in level III increased for neonates 500-1000 g over the study period (1968-1974 Level I 2.71, Level II 0.84 versus 1990-1994 Level I 0.31, Level II 0.35). Infants 1501-2000 g born in 1990-1994 showed less favorable outcome if born in Level III (odds ratio >1) which may be due to severity.	Very low birth weight neonates born in hospitals with level III neonatal services have a higher probability of surviving compared with infants born in hospitals without such services. Case-mix: Level II

Bowes WA 1981 United States	To assess the impact of regionalization of perinatal health care on the neonatal and fetal mortality rates in Colorado.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality (death within the first 28 days of life) Study period: 1971-1974 vs 1975-1978 Comparisons: Level III (3 hospitals) vs Level II (7 hospitals) vs Level I (all others) Study population: Colorado Data source: Bureau of Vital Records of the Colorado State Health Department	No	Neonatal Mortality Rate (per 1000 live births) 1974 Level I: 300 Level II: 346 Level III: 322 1978 Level I: 250 Level II: 196 Level III: 213	The decrease in neonatal mortality appears to be due to improved survival of VLBW infants, which occurred in association with a greater proportion of such infants being delivered in hospitals with level II and level III perinatal services. Case- mix: Level 0

Clarke M 1993 United Kingdom	To evaluate mortality rates as a method of auditing obstetrics and neonatal care after account had been taken of transfer between hospitals during pregnancy and case mix.	Data collection: retrospective Design: observational Primary outcome: perinatal mortality rates Study period: 1978-1987 Comparisons: consultant units vs general practitioner units Study population: Leicestershire Data source: not stated	Yes Demographic: mother <18 and >35 years old, primparity or parity of >3, manual vs non-manual work, <158cm tall, antenatal care after 18 weeks gestation, diabetes, smoker, Asian origin, history of infertility, husband in social class 3-manual, 4, or 5, no general practitioner on the obstetrics list, birth weight, urinary tract infection, pre-eclamptic toxaemia, antepartum hemorrhage, >7 days inpatient care, no antenatal care	Adjusted Perinatal Mortality Rates (per 1000 deliveries) Consultant units: 10.1-12.4 General practitioner units: 3.8	Perinatal mortality rates by place of delivery showed substantial differences between general practitioner units and consultant units because of the selective referral of high risk women from general practitioners units to consultant units. Case-mix: Level I

Cordero L 1988 United States	To describe the five-year experience of a regional perinatal program in a defined geographic area using the outcome of very-low birth weight infants as a marker of success (1979-1983).	Data collection: retrospective Design: observational Primary outcome: neonatal mortality (<28 days) Study period: 1979-1983 Comparisons: Level III (regional perinatal centre) vs Level II (community hospitals) vs Level I (rural hospitals and one community hospital) Study population: 33 counties in central and southeast Ohio Data source: linked birth and death certificates from Ohio Department of Health	Yes Stage/severity of illness: birth weight	Neonatal Mortality (%) Level I 500-750g: 83 751-1000g: 60 1001-1250g: 26 total: 54 Level II 500-750g: 88 751-1000g: 52 1001-1250g: 18 total: 52 Level III 500-750g: 68 751-1000g: 43 1001-1250g: 13 total: 36	Higher mortality at Level I and II centres underscores the importance of place of birth on survival of the VLBW infant. Case-mix: Level II
Field 1991 United Kingdom	To examine how differing approaches to neonatal care, decided electively, affected survival after premature delivery in one region within the United Kingdom.	Data collection: prospective Design: observational Primary outcome: survival to discharge or death of infants born at ≤28 weeks gestation Study period: Feb 1, 1987- Jan 31, 1988 Comparisons: large units (≥500 days of ventilation annually) vs small units (<500 days of ventilation) Study population: Trent region (UK) Data source: collected at hospital admission by investigators	Yes Stage/severity of illness: scoring system based on birth weight, gestation, respiratory distress at birth, cephalic or breech presentation, Apgar scores at 1 and 5 minutes, and multiple pregnancy	Relative Odds of Death (small units compared to large units by gestational age [weeks]): >32 but ≤34: 0.97 (0.25,3.70) >30 but ≤32: 1.26 (0.50,3.23) >28 but ≤30: 0.34 (0.07,1.61) ≤28: 3.98 (1.55,10.18)	These results support the policy that there should be a more centralized neonatal service for those infants at or below 28 weeks' gestation. Case-mix: Level II

Gortmaker S 1985 United States	To estimate differentials in survival among very low-birth weight infants according to hospital of birth, and to determine importance of birth at high-technology centres versus birth at other urban or rural hospitals.	Data collection: retrospective Design: observational Primary outcome: infant mortality rate (death <28 days) Study period: 1978-1979 Comparisons: Level III (NICU) vs other urban (Level II and I within a standard metropolitan statistical area) vs rural (Level II and I outside a standard metropolitan statistical area) Study population: Louisiana, Ohio, Tennessee, and Washington Data source: linked birth and death certificates	Yes Demographic: race, length of gestation period, and multiple vs single birth Stage/severity of illness: birth weight	In general, infants born in level III centres had the best survival, while infants born in the rural hospitals experienced the worst survival. Example: 750-1000 g (white infants, early neonatal period) Level III: reference Other urban: 1.36 ($p<0.01$) Rural: 1.45 ($p<0.01$) 1000-1500 g (white infants, early neonatal period) Level III: reference Other urban: 1.51 ($p<0.001$) Rural: 1.68 ($p<0.001$)	There are significant differentials in the survival of infants of similar race and birth weight, depending on their hospital of birth; very low birth weight infants benefit greatly from being born in a Level III perinatal centre. Case-mix: Level II

Lagoe RJ 1986 United States	To evaluate selected aspects of neonatal care utilization by DRG (diagnostic related groups) in four federally defined Health Service Areas of Upstate New York with regional neonatal programs.	Data collection: retrospective Design: observational Primary outcome: neonatal care utilization Study period: 1983 Comparisons: Level III (regional neonatal centres) vs Level II (area neonatal units) vs Level I (community neonatal units) Study population: Upstate New York Data source: New York's Statewide Planning and Research Corporation System and the State's Appropriateness Review of neonatal care services	No	Mean Length of Stay Neonates, died or transferred Level I: 1.85 Level II: 5.04 Level III: 14.90 Extreme immaturity, neonates Level I: 9.38 Level II: 23.26 Level III: 45.77 Prematurity with major problems Level I: 9.38 Level II: 14.05 Level III: 23.46	The diagnosis -related groups of neonates, died or transferred; extreme immaturity, neonates; and prematurity with major problems, generate substantially different utilization patterns in Level III, Level II, and Level I facilities Case-mix: Level 0
LeFevre M 1992 United States	To examine the relationship between neonatal mortality and the level of perinatal services present in the hospital of birth.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality Study period: 1980-1984 Comparisons: Level III (NICU and diagnose and treat severe respiratory distress syndrome requiring mechanical ventilation) vs Level IIB (NICU) vs Level IIA (no NICU) vs Level IB (= 500 deliveries/yr) vs Level 1A (<500 deliveries/yr) Study population: Missouri Data source: birth certificates	Yes Demographic: race Stage/severity of illness: birth weight	The mortality rate was higher for black and white babies born at Level I-A hospitals who weighed less than 2250 g compared to those born at Level III hospitals. There were no other statistically significant differences. For white babies less than 1500 g born at Level III, there was a trend for lower mortality.	The results of this study support the safety of facilities with lower levels of care for delivery of normal birth weight infants and the need for continued centralized delivery of higher levels of care for high-risk patients. Case-mix: Level II

Linder-Pelz S 1990 Australia	To compare the obstetric outcomes for women admitted to the Birth Centre at Royal Hospital for Women in Sydney with outcomes for women admitted to the conventional labour ward.	Data collection: retrospective Design: observational Primary outcome: obstetric outcomes Study period: not stated Comparisons: conventional labour ward vs birth centre at Royal Hospital for Women Study population: women delivering at Birth Centre at Royal Hospital for Women and those at a "conventional labour ward" Data source: medical records	Yes Co-morbidities: prenatal and intrapartum risk	Neonatal Death Rate (per 1000 live births) Birth centre: 4.5 Labour ward: 5.9	The comparison of results for admissions to the two settings at Royal Hospital for Women showed that the outcomes for the birth centre were at least as good as those for the conventional labour ward, after differences in prenatal and intrapartum risk were taken into consideration. Case-mix: Level II
Mayfield JA 1990 United States	To establish more clearly the relationship of hospital obstetrical volume and nursery level of care to outcome by using the perinatal mortality rate as the outcome measure of quality.	Data collection: retrospective Design: observational Primary outcome: perinatal mortality (fetal deaths >1000g and all live births >1000g which died =7 days) Study period: 1980-1983 Comparisons: Level III (neonatal intensive care unit) vs Level II (intermediate) vs Level I (normal newborn facility) Study population: Washington state Data source: Washington State linked birth-death records and fetal death records	Yes Stage/severity of illness: birth weight Other: maternal risk factors	Birth Weight Specific Perinatal Mortality Rates (per 1000 births) for Infants 1001-1500 g / 1501-2000 g: Level I: 459 / 170 Level II: 358 / 150 Level III: 175 / 68 For infants >2000 g, level of care had little influence on survival rates. Loglinear analyses showed that nursery level had a stronger influence on mortality than obstetrical volume.	The effect of hospital technology upon perinatal mortality appears to be greater than the effect of volume, but this benefit is enjoyed mainly by the low birth weight, high-risk infant. Case-mix: Level II

Paneth N 1987 United States	To present findings for the entire population of singleton hospital births in New York City during the period from 1976 to 1978, and to consider the implication of these combined findings for perinatal health care provision in urban areas.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality (within the first 28 days) Study period: 1976-1978 Comparisons: Level III (NICU) vs Level II (intermediate care) vs Level I (no special facilities to care for sick infants) Study population: New York City Data source: medical records	Yes Demographic: gestational age (in addit ion, analysis of full-term infants with normal birth weights controlled for maternal education, race, sex, maternal age, and parity) Stage/severity of illness: birth weight	Birth Weight Standardized Neonatal Mortality Rates (per 1000 live births) Level I: 9.49 Level II: 9.79 Level III: 9.12 Mortality was significantly higher for level I and level II births combined vs level III for: gestation <37 weeks (all weight categories) and gestation =37 wks (birth weight 1251 -2250 g).	The evidence suggests that when labor is established before 37 weeks' gestation, or when intrauterine growth retardation occurs, delivery in a tertiary care setting carries reduced risk. Case-mix: Level II

Roder D 1993 Australia	To investigate differences by birth weight in risk of perinatal death between level III hospitals (which provide care for high risk pregnancies and neonatal intensive care) and other hospitals in South Australia, using perinatal data for the 1985-1990 period.	Data collection: retrospective Design: observational Primary outcome: perinatal mortality (live and still births =400g or 20 weeks gestation that die in first 28 days) Study period: 1985-1990 Comparisons: Level III (hospitals with facilities for high risk pregnancies and NICU) vs Level II (teaching hospitals with special neonatal facilities) vs Level I (county hospitals of >100 births per year and smaller county hospitals) Study population: South Australia Data source: data collected by midwives and neonatal nurses verified by Registrar of Births, Deaths, and Marriages	Yes Demographic: gestational age at birth, mother's age and race, parity Stage/severity of illness: birth weight Co-morbidities: congenital abnormalities	Relative Odds of Perinatal Death for Babies Born in a Non-Level III Hospital (adjusted for important risk factors): <1000: 10.04 (4.94,20.38) 1000-1499: 5.08 (3.18,8.10) 1500-1999: 2.11 (1.38,3.23) 2500-2999: 0.56 (0.40,0.79) Same analyses restricted to babies without congenital abnormality: <1000: 9.24 (4.50,18.99) 1000-1499: 5.35 (3.28,8.72) 1500-1999: 2.30 (1.40,3.79) 2500-2999: 0.65 (0.44,0.98)	After adjustment for birth weight, risk estimates changed such that a lower overall risk of death was evident for babies born in level III hospitals. Case-mix: Level III

Rosenblatt RA 1985 New Zealand	To assess whether the low volume of deliveries in the many remaining small hospitals affects perinatal mortality, in the context of a regionalized system of care.	Data collection: retrospective Design: observational Primary outcome: perinatal mortality (<7days of age) Study period: 1978-1981 Comparisons: Level III (tertiary care units for obstetrics and perinatal care) vs Level II (subregional referral hospital) vs Level I (general practitioners working with midwives) Study population: New Zealand Data source: National Health Statistics Centre of New Zealand	Yes Stage/severity of illness: birth weight	Perinatal Mortality Rate <= 1500g Level I: 547.8 Level II: 581.5 Level III: 408.7 Level II vs Level III: p<0.01 1500-2499g Level I: 46.9 Level II: 71.8 Level III: 62.2 Level I vs Level II: p<0.01 >= 2500g Level I: 3.2 Level II: 5.3 Level III: 6.0 Level I vs Level II: p<0.01 Level I vs Level III: p<0.01	Level I maternity hospitals had lower birth-weight specific perinatal mortality rates in all but the lowest birth-weight categories than the better equipped hospitals to which they refer. Case-mix: Level II

Rosenblatt RA 1988 United States	To evaluate the extent to which the regionalization of perinatal care in Washington state has succeeded in concentrating high-risk pregnancies in technologically appropriate referral centres and in reducing differences in neonatal outcome among hospitals.	Data collection: retrospective Design: observational Primary outcome: standardized mortality ratios (neonatal deaths 0-27 days) Study period: 1980-1983 Comparisons Level III (sophisticated obstetric and neonatal capabilities) vs Level II (varied levels of respiratory support available, but all were able to resuscitate, treat sepsis, do exchange transfusions, and manage intra-arterial lines) vs Level I (uncomplicated obstetric and neonatal care) Study population: Washington state Data source: Washington State linked birth and infant death records	Yes Stage/severity of illness: birth weight	Standardized Neonatal Mortality Ratio Level I: 1.02 (0.93, 1.12) Level II: 1.03 (0.93, 1.14) Level III: 0.97 (0.90, 1.05)	The standardized mortality ratios for each level of care are near unit; none of the observed differences are statistically significant. Level III hospitals had significantly lower mortality rates for infants weighing less than 1500g. Case-mix: Level II

Rosenblatt RA 1996 United States and United Kingdom	To compare perinatal regionalization and neonatal mortality in Wales and Washington State.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality Study period: 1989-1990 Comparisons: Wales (less regionalized) vs Washington (regionalized) Study population: Washington State and Wales Data source: Wales: Office of Population Censuses and Surveys Washington: Department of Health of the State of Washington	Yes Demographic: birth weight	Crude Neonatal Mortality Rates >1000 grams Wales: 2.47/1000 Washington: 2.32/1000	To the extent that birth weight-specific neonatal mortality rates reflect the quality of care, the two systems (one more regionalized than the other) achieve comparable results for babies weighing >999 grams at birth. Case-mix: Level II
Sangala V 1990 United Kingdom	To determine the perinatal mortality rate among normally formed, singleton babies with birth weights =2500g in Bath health district based on the intended place of delivery at the time of onset of labour or at the time of diagnosis of intrauterine death.	Data collection: prospective Design: observational Primary outcome: delivery outcomes Study period: Oct 1984-Dec 1987 Comparisons: consultant unit vs integrated general practitioner vs isolated general practitioner unit Study population: Bath health district Data source: medical records	Yes Demographic: parity	Perinatal Mortality Rate Consultant unit: 2.8 (1.6, 5.0) Integrated general practitioner unit: 0(0, 2.4) Isolated general practitioner unit: 4.8 (2.9, 6.7)	It is inappropriate for the adjusted perinatal mortality rate to be almost two times higher in the isolated general practitioner units than the consultant unit, which was taking the high-risk deliveries; the authors state that this finding is alarming and requires explanation. Case-mix: Level I

Verloove-Vanhorick SP 1988 Netherlands	Dutch data concerning very preterm or very low birth weight infants born in the Netherlands in 1983 were used to evaluate the influence of place of birth on mortality.	Data collection: retrospective Design: observational Primary outcome: in-hospital mortality Study period: 1983 Comparisons: Level III (perinatal ICU in university hospitals) vs Level II (limited facilities for special care) vs Level I (little or no facilities for special care) Study population: Netherlands Data source: Project on Preterm and Small for Gestational Age	Yes Demographic: gestational age, sex Stage/severity of illness: birth weight Co-morbidities: multiple pregnancies Numerous other perinatal variables were controlled for	OR for In-hospital Death (adjusted for gestational age, birth weight, sex, and multiple birth): Level III: reference Level II: 1.56 (1.04, 2.33) Level I: 1.62 (1.12, 2.34) OR for In-hospital Death (adjusted for 22 perinatal variables): Level III: reference Level II: 1.90 (1.1,3.2) Level I: 1.80 (1.1, 3.0)	The adjusted mortality odds were significantly higher in level I and level II compared with level III hospitals. Case-mix: Level III
Viitsainen K 1994 Finland	To study whether hospitals of different levels are equally safe places to give birth in a regionalized system of care.	Data collection: retrospective Design: observational Primary outcome: low birth weight (<2500g), prematurely (<37 weeks), crude and birth-weight specific perinatal mortality rates, <7 one minute Apgar score, newborns needing transfer to newborn surveillance, special care unit, or NICU Study period: 1989-1990 Comparisons: Level III (NICU) vs Level II (adjacent newborns special care units) vs Level IB (local hospitals and one big community hospital) vs Level IA (community hospitals	Yes Demographic: maternal age, marital status, education, previous pregnancies, previous stillbirths, catchment area Stage/severity of illness: birth weight	Crude Perinatal Mortality OR (Adjusted for maternal background characteristics) Level III: reference Level II: 0.92 (0.72, 1.16) Level IB: 1.02 (0.76, 1.36) Level IA: 0.58 (0.18, 1.86)	No statistically significant differences were found in crude or birth-weight specific perinatal mortality rates between the catchment areas, nor did other outcomes studied favor tertiary care compared with other levels of care in the area-based analysis. Authors conclude that small local hospitals are safe in a regionalized system of birth care. Case-mix: Level II

		with maternity care for low risk pregnancies with no special care units and in remote areas) Study population: Finland Data source: Finnish Medical Birth Registry			
Yeast JD 1998 United States	To study changing patterns of low-birth weight outcome over the past decade as deregionalization of perinatal care has occurred.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality Study period: 1982-1986 and 1990-1994 Comparisons: Level III (>2000 deliveries/yr and/or an NICU average daily census of >10, intrapartum and neonatal intensive care) vs Level II (1001-1999 births/yr stabilization and transfer of complicated obstetric cases) vs Level I (=1000 births/yr, uncomplicated maternity and neonatal care) Study population: Missouri Data source: linked birth and death certificates	Yes Demographic: race, number born (singleton vs multifetal pregnancy) Stage/severity of illness: birth weight	OR of Neonatal Mortality for ELBW (500-999 g): Level III (1990-1994): reference Level III (1982-1986): 2.19* Level II: 2.55* (1990-1994); 2.77* (1982-1996) Level I: 4.66* (1990-1994); 4.83* (1982-1986) OR of Neonatal Mortality for VLBW (1000-1499 g): Level III (1990-1994): reference Level III (1982-1986): 1.83 Level II: 2.28* (1990-1994); 2.57* (1982-1986) Level I: 2.57* (1990-1994); 3.95* (1982-1996) (* significant at p=0.05)	Changing patterns of perinatal regionalization have not improved outcomes for inborn infants <1500g except in level III centres. Case-mix: Level II

Neonatal: Volume or Regionalization

Author / Year/ Country	Objective	Study Characteristics	Case Mix	Results	Author's Conclusions
Bakketeig LS 1978 Norway	The focus of the study is on differences in perinatal mortality of the offspring of mothers living in areas where different obstetric services were available at the time of birth.	Data collection: retrospective Design: observational Primary outcome: perinatal, fetal, and early neonatal (<7 days) mortality rates Study period: 1967-1973 Comparisons: counties grouped by average obstetric score: high (>140) vs medium (101 - 140) vs low (=100) Study population: Norway Data source: National Health Services of Norway and Medical Birth Registry of Norway	Yes Demographic: parity, maternal age, marital status Stage/severity of illness: birth weight	Overall Perinatal Mortality Rates by Level of Obstetric Aid: High: 21.0 Medium: 20.7 Low: 22.3 Differences between low vs medium and low vs high are significant ($p<0.01$) with adjustment for birth weight.	Overall perinatal mortality rate is highest in the group with the lowest level of obstetric aid. Results are consistent for early neonatal mortality and fetal mortality. Case-mix: Level II
Callon HF 1975 United States	To discusses regionalization of perinatal care in Wisconsin.	Data collection: retrospective Design: observational Primary outcome: not stated Study period: 1960-1973 Comparisons: pre-regionalization vs post-regionalization-regionalization occurred in 1971 Study population: Wisconsin Data source: Wisconsin Bureau of Vital Statistics	No	Perinatal death rates have decreased from 17/1000 live births in 1960 to approximately 10/1000 in 1973.	Wisconsin perinatal death rates have decreased since the beginning of the program. Case-mix: Level 0

Campbell 1984 United Kingdom	To determine possible reasons for variations in admission rates observed in the neonatal care units of the East Anglian Regional Health Authority.	Data collection: retrospective Design: observational Primary outcome: admission rates Study period: not stated Comparisons: not stated Study population: Anglian Regional Health Authority Data source: not stated	No	Among the variables examined, size of the hospital was the strongest correlate of NICU admission rates (Kendall's t – 0.90, p<0.002).	Admission rates to neonatal care units are highest in hospitals with fewer deliveries, those with no registrar coverage, in hospitals that are not attached or affiliated to a medical university school and in hospitals with the more liberal admission policy. Case-mix: Level 0
Deutchman ME 1995 United States	To compare obstetrician and family physician patient population demographics, obstetric outcomes, delivery methods, and medical risk factors.	Data collection: retrospective Design: observational Primary outcome: not stated Study period: May 1991 - Dec 1992 Comparisons: 16 obstetricians vs 28 family practice faculty and residents Study population: urban community hospital Data source: birth certificates and fetal death certificates	Yes Demographic: Goodwin index (base-line risk including age, parity, preexisting medical conditions, pregnancy complications, gestational age, maternal age <16 years, diabetes, hypertension, bleeding)	No significant difference was evident in perinatal morbidity or mortality between family physicians and obstetricians (despite difference in methods of delivery). Example: median Apgar scores were 8 and 9 at 1 and 5 minutes, respectively.	The findings show that family practice group can provide perinatal care (prenatal and intrapartum) that results in outcomes comparable with or better than those of ob-gyn practicing in the same location. With a medically similar patient population, family physician care resulted in a significantly lower c-section rate, increased vaginal birth after c-section, and equivalent infant outcome. Case-mix: Level I

Fallis G 1988 Canada	To examine birth weight-specific perinatal mortality rates to determine the relative safety of small hospital obstetrics in Canada.	Data collection: retrospective Design: observational Primary outcome: birth-weight specific mortality rates (number of early fetal deaths >20 weeks gestation and weighing >500g plus early neonatal deaths less than 7 days) Study period: 1970-1985 Comparisons: high (=400 deliveries per year) vs medium (101-400) vs low (<100) Study population: Ontario, Newfoundland, Saskatchewan Data source: Hospital Medical Records Institute	Yes Stage/severity of illness: birth weight	Birth weight-specific mortality rates for newborns weighing greater than 2500 g was lower in small hospitals. Ontario =3000 births: 2.6 401-2999: 2.5 <400: 2.0 Saskatchewan =3000: 3.2 401-2999: 1.8 <400: 0.6 Newfoundland =3000: 2.9 401-2999: 3.5 <400: 1.2	The authors have shown that in small hospital obstetrics, birth weight specific mortality for babies weighing greater than 2500 grams is low. Case-mix: Level II
Hein HA 1978 United States	To review the experience of a rural Midwestern state, Iowa, which currently provides maternity care for 42 000 patients per year in 135 hospitals, 82% of which deliver fewer than 500 babies per year.	Data collection: retrospective Design: observational Primary outcome: fetal and neonatal mortality Study period: 1972-1976 Comparisons: >1000 vs 500-999 vs 250-499 vs 100-249 vs 1-99 Study population: Iowa Data source: not stated	No	Neonatal Death Rate ‰ 1972 >1000: 14.15 500-999: 14.18 250-499: 12.94 100-249: 10.56 1-99: 13.80 1976 >1000: 12.26 500-999: 9.34 250-499: 7.04 100-249: 6.11 1-99: 5.85	Small obstetric services can offer acceptable perinatal care when they are part of a regionalized system. Case-mix: Level 0

Hein HA 1980 United States	To present an evaluation of Iowa's system by a review of standard birth and mortality data.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality (first 28 days of life) Study period: 1972, 1978 Comparisons: Level III (tertiary care) vs Level II (regional perinatal centres) vs Level IB (>500 deliveries/year) vs Level IA (<500) Study population: Iowa Data source: Iowa State Health Department	Yes Stage/severity of illness: birth weight	Neonatal Death Rate/1000 Live Births: 1978: 8.3 1972: 13.7	The authors urge the referral of all gravidas in early preterm labor to the tertiary centre, because they feel that fetal maturity may not be adequately assessed in the Level I and, possibly in the level II hospitals. Case-mix: Level II
Hein HA 1981 United States	To describe another method based on neonatal mortality data that can be used to evaluate the outcome of care in a regional perinatal system and provide information about the causes of neonatal death in Iowa and demonstrate how this can be used to determine educational needs within the Iowa regionalized care system.	Data collection: retrospective Design: observational Primary outcome: idealized neonatal mortality rate (within 28 days) Study period: 1978-1979 Comparisons: level III (1 hospital) vs level II (10 hospitals) vs level I (122 hospitals) Study population: Iowa Data source: Iowa State Health Department	Yes Stage/severity of illness: extremely low birth weight (<750g) Co-morbidities: congenital malformations	Excess Death Rate* (per 1000 live births) Level III: 9.2 Level II: 5.5 Level I: 3.9 Excess mortality rate=actual mortality rate-idealized rate, where idealized rate takes into account congenital malformations and ELBW infants	The authors urge the referral of all gravidas in early preterm labor to the tertiary centre, because they feel that fetal maturity may not be adequately assessed in the Level I and, possibly in the level II hospitals. Case-mix: Level II

Hein HA 1986 United States	The purpose of this paper is to describe the Iowa regional perinatal care system and to evaluate the effects of this system on perinatal outcome in the state.	Data collection: retrospective Design: observational Primary outcome: mortality Study period: 1982 Comparisons: Level III vs Level II vs Level I (<500 deliveries/yr) Study population: Iowa Data source: Iowa State Health Department	Yes Stage/severity of illness: birth weight	Neonatal Mortality Rate Level III: 14.4 (1982) vs 23.3 (1972) Level II: 14.6 (1982) vs 6.6 (1972) Level I (>500 deliveries/yr): 14.4 (1982) vs 5.2 (1972) Level I (<500 deliveries/yr): 12.8 (1982) vs 3.0 (1972)	Birth and mortality data comparing the year immediately prior to the development of a regionalized system of care in Iowa with the most current year having complete results suggest that pregnancy outcome has improved and that a stratified system of care is functioning in the state. Case-mix: Level II
Hein HA 1986 United States	The purpose is to describe the Iowa regional perinatal care system and to offer evaluation of the effects of this system on perinatal outcome in the state.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality in first 27 days Study period: 1972 and 1982 Comparisons: 1972 (last year prior to regionalization) vs 1982 (post-regionalization) Study population: Iowa Data source: Iowa State Health Department	Yes Stage/severity of illness: birth weight	Overall difference in neonatal mortality rates in 1972 and 1982 is significant: χ^2=120.84 ($p<0.001$) The reduction in neonatal mortality rates was much more dramatic in lighter birth-weight classes. Similar trends observed for fetal mortality rates.	Birth and mortality data comparing the year immediately prior to the development of a regionalized system of care with the most current year show lower birth-weigh-specific mortality in all categories except fetal deaths weighing <1001 grams. Case-mix: Level II

Hein HA 1999 United States	To provide a modern update of the causes of neonatal deaths to help focus prevention measures.	Data collection: retrospective Design: observational Primary outcome: idealized neonatal mortality rate (within 28 days) Study period: 1982-1983 and 1995-1996 Comparisons: level III (2 hospitals) vs level II (10 hospitals) vs level I (95 hospitals) Study population: Iowa Data source: Iowa Department of Public Health	Yes Stage/severity of illness: extremely low birth weight (<700g) Co-morbidities: congenital malformations	Excess Mortality Rate* (per 1000 live births) Level III: before 9.7, after 1.8 Level II: before 2.4, after 0.7 Level I: before 1.5, after 0.8 Excess mortality rate=actual mortality rate-idealized rate, where idealized rate takes into account congenital malformations and ELBW infants	The overall reduction in the number of deaths statewide and the movement of deaths to the tertiary centres are consistent with the goals of a regionalized system of perinatal health care. Case-mix: Level II
Hoekstra R 1981 United States	To describe a regionalized community perinatal program and the effect that this program has had on the intermediate care hospital involved.	Data collection: retrospective Design: observational Primary outcome: admissions, length of stay, neonatal and perinatal mortality Study period: 1978 and 1980 Comparisons: 1978 (pre-program implementation) vs 1980 (post-program) Study population: Minnesota or Minneapolis Data source: natality statistics at Fairview-Southdale; statistics from Children's Perinatal Centre	Yes Variables not reported	Neonatal Mortality at Fairview-Southdale 1980: 0/1000 1978: 4.7/1000 live births	Neonatal mortality and morbidity decreased significantly after initiation of a community based regionalized perinatal program. Case-mix: Level I or more

International Neonatal Network, Scottish Neonatal Consultations, Nurses Collaborative Study Group 2000 United Kingdom	To compare perinatal care for all VLBW infants in Scotland and Australia.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality (<27 days) Study period: 1993-1994 Comparisons: Australia with 6 NICU hospitals vs Scotland with tertiary and non-tertiary hospitals Study population: all live born infants 500-1499 g in Scotland and Australia Data source: statutory sources of neonatal mortality statistics	Yes: Stage/severity of illness: birth weight	Neonatal mortality for infants 500-1499 g was higher in Scotland than Australia (RR 1.22; 1.08,1.39).	Perinatal outcome was better in Australia (6 NICU hospitals) compared to Scotland (with tertiary and non-tertiary hospitals). These results may be due to differences in the organization and implementation of neonatal care. Case mix: Level II
Klein MC 2002 Canada	To determine if the practice-volume relations that have been shown in other fields of medical practice also exist in maternity care practice by family doctors.	Data collection: retrospective Design: observational Primary outcome: maternal morbidity, 5-minute Apgar, and admission to NICU or special care unit (SCU) Study period: April 1997-August 1998 Comparisons: high (>25 deliveries each year) vs medium (12-24) vs low (<12) Study population: British Columbia's Women's Hospital and Health Centre Data source: hospital records for the use of research and quality assurance	Yes Demographic: parity, ethnicity, lone parent status, maternal age, gestational age Stage/severity of illness: head circumference, birth weight Other: pregnancy induced hypertension, gestational diabetes (type 1 and 2)	Adjusted OR of 5 Minute Apgar Score <7 >25: 0.908 (0.540, 1.524) 12-24: 0.652 (0.339, 1.251) <12: reference Adjusted OR of Admission to NICU or SCU >25: 0.849 (0.610,1.181) 12-24: 0.862 (0.584,1.274) <12: reference Complex Maternal Morbidity >25: 0.960 (0.743,1.242) 12-24: 1.137 (0.845,1.529) <12: reference	Family physicians' delivery volumes were not associated with adverse outcomes for mothers or newborns. Case-mix: Level II

LeFevre ML 1992 United States	To determine the relationship between physician volume and perinatal outcome as measured by neonatal and perinatal mortality.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality (live birth that die within <28days) and perinatal mortality (neonatal and fetal deaths) Study period: 1984-1987 Comparisons: >400 deliveries in 4 years vs 301-400 vs 201-300 vs 101-200 vs =100 Study population: Missouri Data source: fetal death certificates linked to birth certificates	Yes Demographic: race, mother's age, education, and smoking	Adjusted Perinatal Mortality OR 301-400: 0.95 (0.83, 1.09) 201-300: 1.10 (0.95, 1.26) 101-200: 0.98 (0.84, 1.16) =100: 0.86 (0.69, 1.07)	No relationship was found between physician volume and outcome as measured by perinatal neonatal death. Case-mix: Level I
Lumley J 1988 Australia	To assess the perinatal outcome by size of hospital in one Australian State for the years 1982 to 1984.	Data collection: retrospective Design: observational Primary outcome: perinatal death (within 28 days) Study period: 1982-1984 Comparisons: =2000 births vs 1000-1999 vs 100-999 vs 50-99 vs 20-49 vs <20 births Study population: Victoria state Data source: fetal death certificates linked to birth certificates	Yes: Stage/severity of illness: birth weight	Birth Weight Standardized Perinatal Mortality Ratios (excluding lethal malformations) =2000: 92 (86, 99) 1000-1999: 99 (89, 110) 100-999: 114 (103, 187) 50-99: 120 (80-168) 20-49: 128 (85, 179) <20: 102 (48, 177)	Effective regionalization need not involve the closure of small maternity units on the grounds of safety. Case-mix: Level II

McCormick 1985 United States	This report summarizes the evaluation of a national demonstration program of such regionalization that was funded by the Robert Wood Johnson Foundation (RWJF) in 1975.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality (<27 days) Study period: not stated Comparisons: eight regions received funding to impalement regionalization vs eight control Study population: 8 regions that received funding (RWJF) were matched to 8 control regions Data source: birth and matched infant death and birth certificates obtained from state and local health offices	Yes Stage/severity of illness: birth weight	Neonatal Mortality Rates 1970-1971 RWJF: 11.53 Control: 13.12 1974-1975 RWJF: 9.51 Control: 11.05 1978-1979 RWJF: 7.66 Control: 8.24	The centralization of high-risk deliveries appeared so widespread that the special effect of the RWJF program could not be detected. Case-mix: Level II

Muster D 1999 Norway	To examine risk of neonatal death after low risk pregnancies in relation to size of delivery units.	Data collection: retrospective Design: observational Primary outcome: neonatal mortality Study period: 1972-1995 Comparisons: >3000 vs 2001-3000 vs 1001-2000 vs 501-1000 vs 101-500 vs =100 deliveries Study population: Norway Data source: Medical Birth Registry of Norway and Statistics Norway	Yes Demographic: maternal age (<18 or >40) Stage/severity of illness: congenital malformations, breeched, transverse position, placenta previa, maternal disease before or during pregnancy, =4 previous births, gestational age of >42 weeks, hydrmnios or oligohydramnios, premature rupture of membranes	RR of Neonatal Death in Singleton Infants =2500g >3000: 1.7 (1.4, 2.0) 2001-3000: 1.0 reference 1001-2000: 1.5 (1.3, 1.8) 501-1000: 1.5 (1.3, 1.8) 101-500: 1.7 (1.4, 2.0) =100: 2.1 (1.6, 2.8)	Around 2000 to 3000 annual births are needed to reduce the risk of neonatal deaths after low-risk deliveries. Case-mix: Level II

Moster D 2001 Norway	To compare neonatal mortality in geographical areas where most deliveries occur in large hospitals with areas where a larger proportion of deliveries occur in small maternity units.	Data collection: retrospective Design: observational Primary outcome: risk of neonatal death (within one month after birth) Study period: 1967-1996 Comparisons: >3000 vs 2001-3000 vs 1001-2000 vs 501-1000 vs 101-500 vs 100 deliveries Study population: Norway Data source: Statistics Norway (National Population Registry) and Medical Birth Registry of Norway	Yes Demographic: social class differences (unmarried mothers, parity, maternal age), urban-rural differences Stage/severity of illness: birth weight	RR of Neonatal Death (>2500g) >3000: 0.8 (0.7, 0.9) 2001-3000: 0.9 (0.8, 1.0) 1001-2000: 1.0 (1.0, 1.1) 501-1000: 1.1 (0.9, 1.2) 101-500: 1.1 (1.0, 1.2) =100: 1.6 (1.2, 2.1) Controlling for risk factors did no change results.	We observed a small but significantly decreased neonatal mortality in areas where the greatest majority of births occurred in large hospitals. Case-mix: Level II
Peddle LJ 1983 Canada	To review the development of the Nova Scotia Reproductive Care Program and describe results as demonstrated by consumer utilization and perinatal mortality trends throughout the geographic area served.	Data collection: retrospective Design: observational Primary outcome: perinatal mortality (includes stillborn infants of =1000g and neonates of =1000g that die within 7 days) Study period: 1971-1980 Comparisons: Level III (tertiary obstetric units with perinatal intensive care) vs Level II (regional obstetric units) vs Level I (community obstetric units) Study population: Nova Scotia Data source: Nova Scotia Reproductive Care Program	No	Perinatal Mortality (1971-1980) Level I: 23.1 Level II: 29.0 Level III: 47.9	With regionalization of perinatal services, it is possible to reduce the perinatal mortality rate in small community hospitals to levels that approximate those of a sophisticated tertiary care hospital. Case-mix: Level 0

Pickkala P 1985 Finland	The purpose is to report the changes in perinatal mortality during 15 years in one region.	Data collection: retrospective Design: observational Primary outcome: perinatal mortality [fetal deaths and early neonatal death (< 7 days)] Study period: 1968-1982 Comparisons: regionalized: University Central Hospital of Turku has a NIC vs non-regionalized: no NICU in the local eight hospitals; five had a pediatrician and two have a consultant pediatrician since 1978 Study population: catchment area of the University Central Hospital of Turku Data source: death certificates and hospital charts, autopsy records	Yes Stage/severity of illness: lethal congenital malformations, birth weight	Total Annual Perinatal Mortality Rates (per 1000 births) 1968: 17.9 1969: 19.6 1970: 17.8 1971: 15.0 1972: 16.8 1973: 12.5 1974: 13.3 1975: 12.6 1976: 10.9 1977: 13.5 1978: 5.9 1979: 11.2 1980: 9.4 1981: 6.2 1982: 7.0 Trends were similar when birth weight <1000g and lethal congenital malformations excluded.	The authors believe the centralization of obstetric and neonatal services for risk cases, the introduction of modern obstetric and neonatal management, and continuing education of personnel at every level of maternity and neonatal care accounted for the decline in perinatal mortality. Case-mix: Level II

Rozier JC 1980 United States	To review and analyze perinatal mortality at Southeastern General since the initiation of regionalized care in 1975.	Data collection: retrospective Design: observational Primary outcome: perinatal, neonatal, and stillborn mortality Study period: 1970-1978 (regionalization occurred in 1975) Comparisons: 1976=2071 births, 1977=2326, 1978=2196, (1975: transition year) vs 1970=2060, 1971=2089, 1972= 2045, 1973= 1958, 1974=1909 Study population: Southeastern General Hospital in Robeson County Data source: not stated	No	Perinataly Mortality Rate 1978: 21 1977: 23 1976: 19 1975: 24 1974: 27 1973: 26 1972: 26 1971: 32 1970: 35	The full impact of regionalization has yet to be determined. Case-mix: Level 0

Shapiro S 1982 United States	Evaluation of regional systems of care for all high-risk pregnant women and newborn infants were established in eight areas of the United States.	Data collection: retrospective Design: observational Primary outcome: fetal mortality, early and late infant loss, and low birth weights Study period: not stated Comparisons: eight regionalized areas to control regions Study population: programs funded by Robert Wood Johnson Foundation Data source: official health department statistics containing data from live birth, fetal death, and death certificates	Yes Stage/severity of illness: birth weight	The demonstration area did not show a greater decrease in birth weight specific neonatal mortality than all of the comparison areas. No tests of comparison performed; rates are provided for case study area and three comparison areas for three birth weight categories for the years 1978-79, 1974-75, 1970-71.	Improvement in pregnancy outcome could not be attributed directly to the RWJF program in most of the demonstration areas. But, the examination of indicators of regionalization supported the view that regionalization has been playing a significant role in reducing neonatal mortality. Case-mix: Level II
Siegel 1985 United States	To present an overview of the morality and morbidity findings and discuss their policy implications for rural regional perinatal care programs.	Data collection: retrospective Design: quasi-experimental Primary outcome: fetal and neonatal morality rates and short-term obstetric and newborn morbidity Study period: not stated Comparisons: pilot region with centralization vs comparable control region Study population: 5 counties in North Carolina region and a matched control region Data source: vital statistics	Yes Demographic: race Stage/severity of illness: birth weight	Perinatal Mortality Rate Pilot region: 31.0 Control regions: 28.8	Declines were found for both regions, especially for 1501-2500g infants, but no statistically significant differences were observed between regions. While the incidence of prenatal morbidity was the same for both regions, intrapartum and newborn morbidity significantly favored the pilot region. Case-mix: Level II

Siegel 1986 United States	To conduct a controlled, population-based evaluation to test the hypothesis that the regionalized perinatal care program would have the following effect when the study region was compared with the control region: reduction in adverse 1-year developmental, neurological, and maternal-infant attachment outcomes as well as selected physical health measures.	Data collection: retrospective Design: observational Primary outcome: adverse 1-year developmental, neurological, and maternal-infant attachment outcomes Study period: Nov 1978 - Oct 1979 Comparisons: 8 study regions (rural regionalized perinatal care program) vs 8 matched control regions Study population: study and control region in North Carolina Data source: hospital medical records and vital statistic files	Yes Demographic: maternal education, race Stage/severity of illness: birth weight	Receptive language skills based on mothers' reports favored the study region ($p=0.05$); for infants =1500 g, no significant difference in receptive language. No significant differences were found in expressive language scores, cognitive development, motor milestone achievement, abnormal neurologic signs, and developmental reflex maturity.	No significant differences between study and control regions were observed for Bayley mental and motor scores, abnormal neurological signs, and physical health measures. Case-mix: Level II
Taylor J 1989 United States	To examine the impact of the loss of obstetric services at a small rural hospital in Missouri.	Data collection: retrospective Design: observational Primary outcome: not stated Study period: 1984-1987 Comparisons: pre-regionalization (1984-86) vs post-regionalization (1987) Study population: service area of a rural hospital in a Missouri county Data source: birth certificates from Missouri State Department of Health	No	Comparison of perinatal outcomes pre and post-closing of the local hospital showed no significant differences for any outcome variables (numbers not reported).	The closing of the local unit did not appear to have a statistically significant impact on the overall obstetric outcomes of women from the hospital service area. Case-mix: Level 0

Usher R 1977 Canada	To review accurate perinatal mortality data by weight group from (1) the Province of Quebec, (2) the obstetrical services associated with McGill University, and (3) more detailed analysis of deaths at one perinatal referral centre, the Royal Victoria Hospital.	Data collection: retrospective Design: observational Primary outcome: perinatal mortality (fetus weighs >500g and less than 7 days) Study period: 1967-1974 Comparisons: >1000 deliveries each year vs 1-1000 Study population: Quebec Data source: Quebec Perinatal Committee collects information on all births, deaths, and patient transport by birth weight	Yes Stage/severity of illness: birth weight	Mortality Rate >1000 intramural: 15% >1000 referral: 25% >1000 neither: 48% 1-1000: 43%	Regionalized perinatal intensive care can markedly reduce perinatal mortality. Case-mix: Level II
Walker DB 1985 United States	To evaluate the economic outcome of this statewide regionalized neonatal care for very low birth-weight infants using standard methodology for cost-benefit analysis.	Data collection: retrospective Design: observational Primary outcome: economic Study period: 1974-1975 and 1979-1980 Comparisons: 1979-1980 (24591 births in Rhode Island) vs 1974-1975 (22566 births) Study population: Rhode Island Data source: statewide computerized birth-death certificate matching system	Yes Stage/severity of illness: birth weight	When comparing the economic data of 1974-1975 to 1979-1980, the increase in the absolute number of normal survivors since regionalization neonatal care has resulted in benefits surpassing costs by $2 million.	Regionalized neonatal care in the state of Rhode Island has had a positive economic outcome. Case-mix: Level II

Williams RL 1979 United States	An outcome-based index of the effectiveness of perinatal care was empirically tested using 3 441 448 linked birth-death records.	Data collection: retrospective Design: observational Primary outcome: perinatal death [fetal deaths of >20 weeks gestation and neonatal deaths (deaths of live-born infants within 28 days)] Study period: 1960-1973 Comparisons: average annual number of deliveries over the sample period by the number of years that a hospital had a viable maternity service. Study population: California Data source: linked birth-death records supplied by the California Department of Health	Yes Demographic: race, Spanish surname, teenaged mothers (<20 years), White non-Spanish deliveries, ownership of centre, affiliation with medical school, presence of NICU, for-profit vs not for-profit ownership, RN services, specialist-to-generalist ratio, perinatal study committee, urban/rural Stage/severity: Apgar, c-section	The main regression model showed that the effectiveness of perinatal care was higher in larger hospitals ($p=0.0001$) and in hospitals performing relatively more cesarean sections ($p=0.0001$). The risk-adjusted perinatal mortality rate decreased with the size of delivery service until 2,850 births, then increased thereafter.	The medical care delivered to mothers and newborns appears to be at least twice as effective in lowering perinatal mortality for the highest versus lowest volume hospitals, even after accounting for a variety of factors beyond their direct control. Case-mix: Level II

NICU

Author / Year/ Country	Objective	Study Characteristics	Case Mix	Results	Author's Conclusions
Field D 1999 United Kingdom	To compare survival of premature infants, adjusted for disease severity, in different types of neonatal intensive care setting.	Data collection: prospective Design: observational Primary outcome: survival Study period: 1994-1996 Comparisons: 16 neonatal units divided into 5 relatively large units which regularly took outside referrals and 11 smaller units which provided intensive care for a variable proportion of their local population Study population: Trent Health Region Data source: Trent Neonatal Survey database	Yes Stage/severity of illness: CRIB	CRIB Predicted Death (all =32 weeks) Large predicted: 150.5 (135.2, 166.9) Large actual: 160 Small predicted: 105.5 (94.6, 117.6) Small actual: 96	Actual survival rates for infants =32 weeks gestation and for the group of babies =28 weeks gestation fell within the 95% confidence interval of the rate predicted by CRIB for both the larger referral units and the smaller district units. Non-transferred infants =32 wks gestation had better outcome in small units. Taking the study population as a whole, there is evidence that the service is working in an integrated manner. The larger units, appropriately, attract a group of infants who are sicker than those whose care is entirely in the smaller units. Case-mix: Level II

Goodman DC 2002 United States	To determine whether a greater supply of neonatologists or neonatal intensive care beds is associated with lower neonatal mortality.	Data collection: retrospective Design: observational Primary outcome: mortality within first 27 days Study period: 1995 Comparisons: neonatologists: very high (11.6/10000 births) vs high (7.5) vs medium (5.9) vs low (4.3) vs very low (2.7) and intensive care beds: very high (59.3/10000) vs high (40.7) vs medium (32.4) vs low (32.4) vs very low (23.5) Study population: United States Data source: National Centre for Health Statistics	Yes Demographic: sex, single vs multiple birth, maternal age, parity, race, education, extent of prenatal care Stage/severity of illness: birth weight	Neonatal Mortality Adjusted OR Neonatologists Very high: 0.89 (0.83, 0.95) High: 0.91 (0.86, 0.97) Medium: 0.93 (0.88, 0.99) Low: 0.93 (0.88, 0.99) Very low: 1.00 (reference) Intensive care beds Very high: 0.95 (0.89, 1.02) High: (0.93 (0.88, 0.99) Medium: (1.02 (0.96, 1.08) Low: 0.92 (0.86, 0.98) Very low: 1.00 (reference)	There was little difference in the risk of neonatal death between the lowest and highest quintiles of bed supply. The risk was lower in regions with a low supply of neonataolgists that in regions with a very low supply; however, little additional benefit in survival was seen with further increases in supply. Case-mix: Level II

Horbar JD 1988 United States	To determine whether two specific outcomes, survival and survival without supplemental oxygen as of day 28 after birth, varied among NICUs in US, Canada and England.	Data collection: retrospective Design: observational Primary outcome: survival and survival without supplemental oxygen at 28 days Study period: 1983-1984 Comparisons: 11 NICUs with volume ranging from 86-358 live births Study population: voluntary response to request for data from an international sample of neonatologists Data source: data collected from preexisting clinical databases and chart review	No	Crude Number of Infants Surviving NICU A: 69/86 (80%) NICU B: 87/95 (92%) NICU C: 79/97 (81%) NICU D: 87/100 (87%) NICU E: 111/128 (87%) NICU F: 131/151 (87%) NICU G: 141/160 (88%) NICU H: 142/173 (82%) NICU I: 176/210 (84%) NICU J: 180/218 (83%) NICU K: 309/358 (86%)	The differences in outcomes observed at the centres did not correlate with the number of live births; volume of patients treated had no effect on outcome. Case-mix: Level 0
Horbar JD 1997 United States	To determine whether either admission characteristics of infants or specific characteristics of the units such as annual patient volume and the presence of a pediatric residency program could account for observed differences in neonatal mortality rates of very low birth weight infants (501 to 1500 g at birth) among units.	Data collection: retrospective Design: observational Primary outcome: mortality within 28 days Study period: 1991-1992 Comparisons: >113 vs 78-113 vs 47-77 vs <47 (volume quartiles) Study population: Vermont Oxford Network (voluntary collaboration NICUs- only US data used for this study) Data source: Vermont Oxford Network Database	Yes Demographic: race, gender Stage/severity of illness: 1 minute Apgar scores, birth weight Co-morbidity: major birth defect	Risk for Mortality Adjusted OR per 10 additional cases: 1.01 (0.98, 1.04)	Mortality rates for very low birth weight infants (501-1500g) vary among NICUs in the US; annual volume does not explain the variation. Case-mix: Level III

International Neonatal Network, Scottish Neonatal Consultants, Nurse Collaborative Study Group 2000 Scotland/ Australia	To compare outcomes of care in selected neonatal intensive care units (NICU) for very low birth weight (VLBW) or pre-term infants in Scotland and Australia.	Data collection: prospective Design: observational Primary outcome: mortality before hospital discharge Study period: 1993-1994 Comparisons: regionalized area (Australia with 6 tertiary centres) vs non-regionalized (Scotland with 5 tertiary and 3 non-tertiary NICU hospitals) Study population: volunteer sample of Scotland and Australia Data source: medical records	Yes Demographic: gestation Stage/severity of illness: fraction of inspired oxygen, worst base deficit, birth weight Co-morbidity: congenital malformations All variables make up CRIB scores	Standardized Mortality Ratio Australia: reference Scottish tertiary: 1.50 (1.31, 1.69) Scottish non-tertiary: 1.26 (0.85, 1.66)	After adjusting for CRIB and gestation, mortality in Scottish NICUs was significantly higher than in Australia's NICUs. Case-mix: Level III
Paneth N 1982 United States	To examine the neonatal mortality rates of low birth weight infants (501-2250g) in three kinds of hospitals in New York City: those with newborn intensive care units (Level III), those with capabilities for the care of most premature infants (Level II), and those without any special facilities for premature newborns (Level I).	Data collection: retrospective Design: observational Primary outcome: neonatal mortality Study period: 1976-1978 Comparisons: Level III: (regional centres) vs Level II (care for more complicated pregnancies and neonatal illness but refer to Level III for complex procedures vs Level I (normal or mildly ill newborns) Study population: New York City Data source: vital records of New York City	Yes Demographic: gestational age, race, sex Stage/severity of illness: medical risk variables, birth weight	Adjusted Neonatal Mortality OR Level III: 1.0 (reference) Level II: 1.37 (1.19, 1.58) Level I: 1.32 (1.12, 1.55)	The techniques of newborn intensive care can improve survival in small babies. Birth at a level III centre lowers neonatal mortality in low-birth weight infants. Case-mix: Level I

Phibbs CS 1996 United States	To examine the effects of neonatal intensive care unit (NICU) patient volume and the level of NICU care available at the hospital of birth on neonatal mortality.	Data collection: retrospective Design: observational Primary outcome: death within the first 28 days of life (if continuously hospitalized, within the first year of life) Study period: 1990 Comparisons: Level III regional or tertiary NICU (=15 admissions and <15) vs Level II+ expanded intermediate NICU (=15 admissions and <15) vs Level II intermediate NICU (=5 admissions and <5) vs Level I (no NICU) Study population: California Data source: birth-infant death cohort file and the California Office of Statewide Health Planning and Development	Yes Demographic: age, mother and father's years of education, mother and father's age, sex, race/ethnicity of infant, mother's HMO status, mother insurance from Medicade, other government insurance, or self insured Stage/severity of illness: birth weight	OR of Death Level III (=15): 0.62 (0.46, 0.84) Level III (<15): 0.97 (0.71, 1.33) Level II+ (=15): 0.96 (0.73, 1.26) Level II+ (<15): 0.90 (0.71, 1.14) Level II (=5): 0.90 (0.64, 1.25) Level II (<5): 0.90 (0.66, 1.24) Level I: reference	Risk-adjusted neonatal mortality was significantly lower for births that occurred in hospitals with large (average census, >15 patients per day) level III NICUs. Case-mix: Level II

The UK Neonatal Group 2002 United Kingdom	To assess whether patient volume, staffing levels, and workload are associated with risk-adjusted outcomes, and with costs or staff well-being.	Data collection: prospective Design: observational Primary outcome: hospital mortality, mortality of cerebral damage and nosocomial bacteraemia Study period: Mar 1998 to Apr 1999 Comparisons: >57 VLBW infants admitted each year vs 35-57 vs <35 Study population: randomly selected from NICUs in UK Data source: unit admission logs	Yes Demographic: gestation, sex, centile of birth weight for gestational age, mode of delivery, mother's use of antenatal steroids Stage/severity of illness: Apgar at 5min, diagnostic category, infant temperature on admission, surfactant therapy, oxygen therapy, blood gas, birth weight	Adjusted Mortality OR High: reference Medium: 1.12 (0.76, 1.64) Low: 0.97 (0.70, 1.34)	Risk-adjusted mortality and mortality or cerebral damage were unrelated to patient volume or staffing provision; however, nosocomial bacteraemia was less frequent in NICUs with low neonatal consultant provision Case-mix: level II

PICU

Author / Year / Country	Objective	Study Characteristics	Case Mix	Results	Author's Conclusions
DeNicola LK 1992 United States	To report on seven consecutive years of PICU admissions and compare these to a 1998 survey of 96 PICUs and a 1979 survey of 159 PICUs conducted by the Ross Planning Associates.	Data collection: retrospective Design: observational Primary outcome: not stated Study period: 1983-1989 (Florida) and 1988 (national) Comparisons: Jacksonville, Florida (centralized) vs nation (a mix of different models) Study population: Jacksonville, Florida and national data Data source: 1988 survey of 96 PICUs (national) Age groups: not stated	No	Mortality Rates Florida: 3.3% National: 8.2%	A coordinated multihospital program for pediatric intensive care resulted in a lower mortality rate compared to the national average. Case-mix: Level 0
Lopez-Herce J 2000 Spain	To describe the organization of pediatric intensive care units in Spain and the medical assistance provided during 1996.	Data collection: retrospective Design: observational Primary outcome: not stated Study period: 1996 Comparisons: 13-18 beds vs 7-12 vs 1-6 Study population: 31/34 PICUs in Spain Data source: written questionnaire sent to medical directors of 34 Spanish PICUs within or linked to the Spanish	No	Mortality Rate >12 beds: 6.3+/-0.8% 7-12 beds: 5.9+/-2.8% =6 beds: 4.8+/-4.0%,	Mortality increased as the size of the ICU increased, perhaps due to the fact that the larger ICUs are situated in tertiary hospitals and thus admit more serious pathology. Case-mix: Level 0

		public health system Age groups: 0-18 years			
Pearson G 1997 United Kingdom / Australia	To compare the ICU mortality rate in children from the area of the Trent Health Authority, UK, with the rate in children from Victoria, Australia, where pediatric intensive care is highly centralized.	Data collection: retrospective Design: observational Primary outcome: mortality rate Study period: Apr 1, 1994- Mar 31, 1995 Comparisons: centralized pediatric intensive care in Victoria vs dispersed pediatric care in Trent Study population: Trent Health Authority, UK and Victoria, Australia Data source: not reported Age groups: <16 years	Yes Demographic: age Stage/severity of illness: PRISM, PIM, diagnosis, response of pupils to light, base excess, arterial oxygen tension, fractional inspired oxygen concentration, systolic blood pressure at time of admission; admission elective; ventilated within 1 h of admission to ICU Co-morbidities: underlying illness	Adjusted OR for Children in Trent Relative to Victoria PRISM: 2.09 (1.32, 3.29) PIM: 2.09 (1.37, 3.19)	The findings suggest that substantial reductions in mortality could be achieved if every UK child who needed endotracheal intubation for more than 12-24 hours were admitted to one of 12 large specialist pediatric ICUs. Case-mix: Level III

Pearson G 2001 United Kingdom	To compare pediatric intensive care activity and outcome for a defined population of children in 1991 and 1999, during a period of organizational change and centralization of pediatric intensive care.	Data collection: retrospective Design: observational Primary outcome: place and rate of ICU admission, use of mechanical ventilation at admission, mortality, length of stay Study period 1991(before) vs 1999 (after) Comparisons: 510 patients/year (1999) vs 277 patients per year (1991) Study population: hospitals in the West Midlands (UK) Data source: collected by primary researcher (1991) and 1999 data from larger database from all hospitals in the West Midlands region (1999); denominator data from Office for National Statistics Age groups: <15 years	No	Deaths in ICU 1991: 26/277 1999: 32/510 (p=0.051) Length of ICU stay 1991: 103hr 1999: 74hr (p=0.0117)	Overall child mortality rate decreased by 0.34 deaths per 1000 children but was not statistically significant. Length of stay in ICU was significantly shorter post-centralization. Case-mix: Level 0

Pollack MM 1991 United States	To compare outcomes from pediatric intensive care in hospitals with different levels of resources.	Data collection: prospective Design: observational Primary outcome: mortality rates Study period: June 1- Nov 30, 1986 Comparisons: tertiary vs non-tertiary Study population: 16 centres that are members of Pediatric Critical Care Study group: Oregon and southwest Washington- (74 of the 76 acute care hospitals) Data source: post-graduate nurses collected information on identified patients (doesn't state how the patients were identified) Age groups: < 18 years	Yes Demographic: age Stage/severity of illness: Pediatric Risk of Mortality, operative status at the end of the ICU admission day	Adjusted OR for Mortality (by mortality risk category) Tertiary: reference Non-tertiary 0-5% risk: 1.13 (0.12, 10.7) 5-30% risk: 2.23 (0.51, 9.67) >30% risk: 8.11 (1.15, 57.0)	In this study, pediatric intensive care provided in tertiary hospitals was associated with improved survival chances among high-risk patients. Case-mix: Level II

Pollack 1994 United States	To determine the importance of the following care factors previously associated with hospital quality on survival from pediatric intensive care: size of the intensive care unit (ICU), medical school teaching status of the hospital housing the ICU, specialist status (pediatric intensivist), and unit characteristics.	Data collection: prospective Design: observational Primary outcome: mortality Study period: Dec 1989-Jan 1992 Comparisons: =7 hospital beds vs 4-6 hospital beds Study population: USA-hospitals randomly selected from a national survey of all hospitals with PICU in 1989 Data source: medical records Age groups: not stated	Yes Demographic: age Stage/severity of illness Pediatric Risk of Mortality, operative status, emergency or elective status, pre-ICU care area, transport method, pre-ICU cardiac massage, system of disease, cause of disease	No significant associations could be demonstrated between severity-adjusted mortality and ICU size, unit coordination, volume of ICU patients, public or private status of the hospital or children's hospital status (no numbers reported).	For issues such as patient referrals and regionalization, specific services, including pediatric intensive care, may be more important than overall hospital characteristics. Case-mix: Level II
Tilford JM 2000 United States	To examine the relationships between the volume of patient and other unit characteristics on patient outcomes in PICUs.	Data collection: prospective Design: observational Primary outcome: mortality and length of stay Study period: 1993-1994 Comparisons: hospital volumes of 1378 patients, 1246, 1196, 971, 811, 725, 703, 618, 615, 598, 566, 499,427, 384, 222, and 147 Study population: 16 centres that are members of Pediatric Critical Care Study Group Data source: Pediatric Critical Care Study Group Age groups: not stated	Yes Demographic: age Stage/severity of illness: Pediatric Risk of Mortality (PRISM) and Pediatric Overall and Cerebral Performance Scores	100-patient increase in PICU volume (adjusted OR) 0.95 (0.91, 0.99) Reduced length of stay (incidence rate ratio) 0.98 (0.975, 0.985)	The volume of patients in PICUs is inversely related to risk-adjusted mortality and patient length of stay. Case-mix: Level II

Transplant

Author / Year/ Country	Objective	Study Characteristics	Case Mix	Results	Author's Conclusions
Cox K 2001 United States	To determine whether there was a correlation between the number of pediatric liver transplants performed at US transplant centres and mortality rate.	Data collection: retrospective Design: observational Primary outcome: mortality rate Study period: 1994-1998 Comparisons: =15 transplants each year vs 11-15 vs 6-10 vs 1-5 Study population: United States Data source: United Network of Organ Sharing Scientific Registry Age groups: <18 years	Yes Demographic: age	1-Year Post-transplant Mortality Children less than 2 years =15: 16% 11-15: 17% 6-10: 22% 1-5: 26% Children greater than 2 years =15: 14% 11-15: 12% 6-10: 18% 1-5: 12%	Because the volume of pediatric liver transplants at a centre correlates with patient survival, the authors recommend that transplants in children less than <2 years of age be limited to high volume centres that perform over 10 pediatric liver transplants each year. Case-mix: level I

Hosenpud JD 1994 United States	To determine the effect of small transplantation volume on transplant outcome.	Data collection: retrospective Design: observational Primary outcome: mortality Study period: Oct 1987-Dec 1991 Comparisons: =9 transplants/year vs <9 Study population: United States Data source: Scientific Registry of the United Network for Organ Sharing Age groups: <18 years	No (for pediatric data)	Percent Mortality =9: 8.3% <9: 16.9% (p<0.001)	These data demonstrate that the risk of mortality is substantially higher in low-volume cardiac transplant centres. Case-mix: Level 0
Schurman SJ 1999 United States	To analyze centre volume effects using the experience reported to the North American Pediatric Renal Transplant Cooperative Study (NAPRTCS).	Data collection: retrospective Design: observational Primary outcome: graft survival Study period: 1987-1995 Comparisons: >100 transplants/year vs 51-100 vs =50 Study population: centres contributing to the North American Pediatric Transplant Cooperative Study (NAPTCS) database Data source: North American Pediatric Transplant Cooperative Study (NAPTCS) database Age groups: <18 years	Yes Demographic: recipient age, race, transplant year, living vs cadaver donor transplants Stage/severity of illness: prior transplants, prior dialysis, transfusion history, anti T-cell antibody induction	Three-month graft survival decreased with decreasing centre size (e.g., for all transplants, % survival was 90.4, 90.2, 88.4 for high, medium, low volume respectively, p<0.02); however, with inclusion of t-cell antibody induction as a co-variate, difference was no longer significant, p=0.10. Similar trend was seen for 5-year graft survival: without t-cell antibody induction p<0.05, with co-variate p=0.17.	There is superior graft survival in the high-volume centres. However, the results were sensitive to inclusion of t-cell antibody induction as a co-variate. The authors suggest that low volume centres need to identify risk factors influencing these outcomes. Case-mix: Level II

Trauma

Author / Year/ Country	Objective	Study Characteristics	Case Mix	Results	Author's Conclusions
Cooper 1993 United States	To determine concordance between regional outcome and national norms (based on results obtained at pediatric trauma centres) with respect to pediatric injury diagnosis, severity and mortality in a state lacking a well-organized trauma system during a one year period in 1989.	Data collection: retrospective Design: observational Primary outcome: frequency of principal anatomic diagnosis, and the mortality associated with that diagnosis Study period: 1989 (NY) and national data since 1985. Comparisons: New York (state without regionalized trauma) vs national pediatric registry with regionalized trauma Study population: NY State compared to National Pediatric Trauma Registry Data source: discharge abstracts Age groups: <15 years	Yes Stage/severity of illness variables: ISS	Survival was some 10 times greater overall in pediatric trauma centres for patients with brain or internal injuries of moderately great severity (ISS 15-19).	For children with skeletal, brain, or internal injuries of moderately great severity, trauma care provided in a centre properly staffed and equipped to meet the special needs of critically injured children offers a distinct survival advantage. Case-mix: Level II

Cosentino CM 1991 United States	To determine the impact of level I pediatric trauma centre designation on demographics and financial reimbursement in a trauma population admitted to a large metropolitan children's hospital.	Data collection: retrospective Design: observational Primary outcome: financial reimbursement Study period: 1985-1989. Comparisons: one year pre-trauma centre designation vs year 1, 2, and 3 post Study population: Chicago, USA Data source: financial records Age groups: not stated	No	Length of stay was longer during first year post-designation (4.9 days Pre, 6.65 days Post Ye ar 1, P=0.042); length of stay was 5.6 days post year 2 and 4.7 days post year 3; length of stay in post year 3 was significantly decreased compared with post year 1 (p=0.005). Mortality rates were comparable in each group (1.7% pre and 2.0% post year 1). Total hospital charges increased significantly in post year 1; total hospital charge was $5,820 pre and $7,691 post (p=0.034); financial loss per patients was $1,149 pre and $1,795 post year 1 (p=0.055); total hospital charges per patient in year 2 post were $8,082 and $7,276 in year 3 post; financial loss per patients during the period increased significantly (year 2 post: $4,262, p<0.005; year 3 post $2,799, p=0.04	More appropriate use of level I pediatric trauma centres is needed to allow these centres to be maximally efficient and cost-effective. Length of stay was longer during first year post-designation and decreased during post year 2 and 3. Mortality rates were comparable in each group. Total hospital charges increased significantly in post year 1, 2, and 3. Case-mix: Level 0

Konvolinka CW 1995 United States	To present relationships between volume and survival outcome for severely and urgently injured patients admitted to Pennsylvania's trauma centres and discusses the implication of the findings.	Data collection: retrospective Design: observational Primary outcome: survival Study period: 1988-1989 Comparisons: hospital: average yearly number of Pennsylvania Trauma Outcome Study (PTOS) patients; physician: average yearly number of PTOS patients per trauma surgeon Study population: Pennsylvania, USA Data source: Pennsylvania Trauma Outcome Study Age groups: not stated	Yes Demographic: age Stage/severity of illness: Revised Trauma Score	No volume-related variable was a significant contributor to predictions of W, where W was the number of additional (or fewer) survivors, per 100 patients treated, than expected.	All trauma centres had acceptable pediatric survival rates regardless of volume. Case-mix: Level II
Nakayama DK 1992 United States	To characterize the care of injured children in both pediatric and nonpediatric trauma centres. Questions: 1) how do trauma centres with pediatric designation differ from the many centres without special designation? 2) how do centres differ in terms of the referral, transfer, patient age, injury type, and severity? 3) what happens in areas where pediatric trauma centres don't exist? 4) do regional differences affect mortality?	Data collection: retrospective Design: observational Primary outcome: mortality Study period: Oct 1986-Sept 1989 Comparisons: two trauma centres that meet criteria to be a Pediatric Regional Resource Trauma Centres vs urban non-pediatric centres vs rural non-pediatric centres Study population: trauma centres in Pennsylvania Data source: Pennsylvania Trauma System Foundation registry Age groups: <15 years	Yes Stage/severity of illness: ISS (injury severity score), RTS (revised trauma score)	TRISS analysis: No statistically significant differences in survival rate were seen among the three groups of trauma centres in any of the P(s) group, although survival was slightly higher in pediatric centres for the P(s) =0.3 and <0.6.	The crude death rates in pediatric centres and urban nonpediatric centres were similar and significantly lower than rural centres. No significant differences in survival were noted when stratified by the probability of survival based on TRISS methodology. Case-mix: Level II

Potoka DA 2000 Unites States	To conduct a retrospective analysis of 13 351 injured children, aged 0 to 16 years, treated at accredited trauma centres in the state of Pennsylvania between 1993 and 1997.	Data collection: retrospective Design: observational Primary outcome: mortality Study period: 1993-1997 Comparisons: pediatric trauma centre (PTC) vs adult trauma centres with qualifications to treat children (ATC AQ) vs ATC level II vs ATC level I Study population: Pennsylvania Trauma Outcome Study registry Data source: Pennsylvania Trauma Outcome Study registry Age groups: 0-16 years	Yes Demographic: age, sex, mechanism of injury Stage/severity of illness: Injury Severity Score, Revised Trauma Score, and Glasgow Coma Scale. Severe injury= ISS>15	Mortality Pediatric trauma centre: 3.6% Adult trauma centre with qualifications to treat children: 4.3% Level II adult trauma centre: 4.7% Level I adult trauma centre: 8.0%	Injured children treated at a pediatric trauma centre or at adult trauma centres with adult qualifications in Pennsylvania had a better outcome than those treated at adult trauma centres. Case-mix: Level II
Tepas JJ 1998 United States	Five years experience recorded in a multi-institutional pediatric trauma registry was analyzed to define the relationship between case volume and outcome as measured by mortality.	Data collection: retrospective Design: observational Primary outcome: mortality Study period: 1990-1994 Comparisons: high (>1000 patients) vs medium (501-1000) vs low (100-500) Study population: National Pediatric Trauma Registry Data source: National Pediatric Trauma Registry Age groups: not stated	Yes Stage/severity of illness: Pediatric Risk Indicator (PRI) (Glasgow Coma Scale + absolute Pediatric Trauma Score)	PRI 0-1 High volume= 0.09% Medium volume= 0.04% Low volume= 0.11% PRI 1-2 High volume= 4.9% Medium volume= 4.2% Low volume= 5.4% PRI >2 High volume= 41.4% Medium volume= 33.6% Low volume=42.0%	The findings detailed above do not support the hypothesis that increasing patient volume is associated with decreasing mortality. This analysis indicates that increasing volume does not automatically induce increased expertise. Case-mix: Level II

Van Rooyen MJ 1995 United States	To determine if pediatric mortality is predicted by age, presence of head trauma, head trauma with a Glasgow Coma Scale (GCS) score, a low Pediatric Trauma Score (PTS), and transport directly to a pediatric trauma centre.	Data collection: retrospective Design: observational Primary outcome: pediatric mortality Study period: Jan - Oct 1988. Comparisons: Level I hospital ranged from 116-429 Study population: Chicago, USA Data source: Chicago Health Department Age groups: <16 years	No	Hospital Deaths Hospital C: 13/429 Hospital E: 2/267 Hospital B: 7/243 Hospital A: 3/202 Hospital D: 2/140 Hospital F: 2/116	Mortality rates did not differ according to pediatric trauma centre patient volume. Case-mix: Level 0

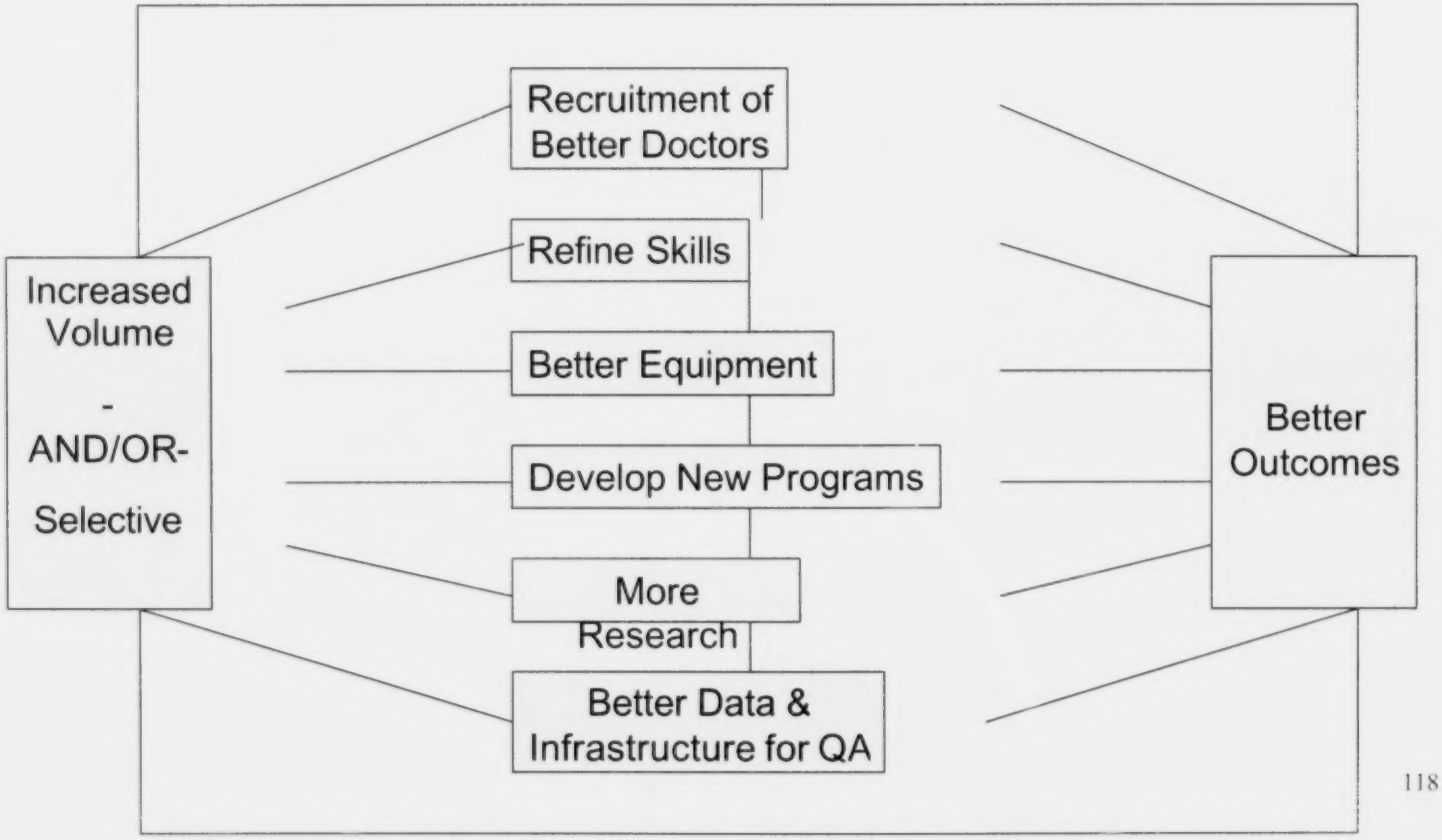
Miscellaneous

Author / Year/ Country	Objective	Study Characteristics	Case Mix	Results	Author's Conclusions
Albright AL 1999 United States	To analyze the outcome a commonly performed pediatric neurosurgical operations: release of tethered cords, and to analyze the neurological outcome data after tethered cord releases.	Data collection: retrospective Design: observational Primary outcome: neurological outcome data Study period: July 1993-June 1998 Comparisons: Surgeon A: 20 years experience vs Surgeon B: 7 years vs Surgeon C: 4 years Study population: not stated Data source: morbidity and mortality records, patient charts and operative records Age groups: not stated-pediatric 0->15 years	No	Post-operative Neurological Deficit Surgeon A (20 yrs experience): 6/49 Surgeon B (7 yrs experience): 0/27 Surgeon C (4 yrs experience): 0/24	It is unclear whether the deficits were associated with other variables, such as neurosurgeons' experience or preoperative neurological status. Case-mix: Level 0
Baumer JH 1997 United States	To investigate the relationship between clinic provision, consultant and nursing caseload, and processes and outcomes of diabetes care in children.	Data collection: retrospective Design: observational Primary outcome: not stated Study period: 1994 Comparisons: specialists treating =40 patients vs non-specialists treating <40 patients Study population: South Western region of England	Yes Demographic: age, duration of diabetes, family history, living arrangements Stage/severity of illness: Townsend deprivation index, insulin dose and	Children under care of tertiary hospital consultants had lowest glycated haemoglobin results ($p<0.001$), spent least time in hospital at diagnosis ($p<0.001$), were most likely to have their heights and weights plotted ($p<0.001$), and to be screened for microalbuminuria ($p<0.001$) and retinopathy ($p=0.035$), had higher admission rates (significance varied by reason for	Children with diabetes under the care of clinicians with a larger caseload appear to have a better clinical outcome than those seen by non-specialists. Study supports national recommendation that consultants providing care for children with diabetes should have a caseload of more than 40 children, and that children be seen in a designated diabetic clinic. Case-mix: Level II

		Data source: members of the diabetes team completed a questionnaire for each child Age groups: not stated	number of injections per day	admission), lower clinic attendance rates ($p<0.001$), and fewer dietary consultations ($p<0.001$). Differences remained significant after adjustment for significant confounders.	
Bratton SL 2001 United States	To determine whether the risk of operative management of children with intussusceptions varies by hospital pediatric caseload.	Data collection: retrospective Design: observational Primary outcome: not stated Study period: 1987-1996 Comparisons: hospitals with 0-3000 annual pediatric admissions vs 3001-10000 vs >10 000 Study population: Washington state Data source: "computerized information" – no other details provided Age groups: <15 years	Yes Demographic: age Co-morbidities: coexisting conditions (other potential confounders not specified)	Adjusted OR for Operative Care >10 000: reference 3000-10 000: 5.3 (3.3, 8.6) <3000: 3.3 (2.0, 5.4) Adjusted OR for Bowel Resection >10 000: reference 3000-10 000: (0.9 (05, 1.7) <3000: 1.0 (0.5, 2.0)	The authors found that children with intussusceptions who received care in a large pediatric hospital had a lower risk of requiring an operative intervention, decreased LOS, and lower hospital charges compared with children treated in smaller pediatric hospitals. Case-mix: Level II
Snow BW 1996 United States	To determine if pediatric surgical subspecialization provides cost-effective, high-quality pediatric patient care.	Data collection: retrospective Design: observational Primary outcome: costs Study period: 1990-1993 Comparisons: 8 general urologists vs 2 pediatric urologists Study population: Primary Children's Medical Centre (University of Utah School of Medicine) Data source: patient records Age groups: 0.25-18 years	No	Mean Length of Stay (days) Pediatric urology: 2.73 General urology: 3.52 ($p<0.00002$) Mean Hospital Charge Pediatric urology: $4645 General urology: $5740 ($p<0.000002$)	General urology patients' mean LOS and hospital charges were significantly less than pediatric urology patients. Rate of complications was lower in pediatric urology patients but not statistically significant. Case-mix: Level 0

Witt PD 1998 United States	To explore the concept of an operative learning curve for different surgeons with respect to palatoplasty.	Data collection: retrospective Design: observational Primary outcome: occurrence of secondary palatal management Study period: 1978-1990 Comparisons: 9 surgeons Study population: computerized data system Data source: computerized data system Age groups: not stated	No	There was a strong relationship between experience (number of procedures performed) and success (no need for secondary palatal management) (linear regression, r^2=0.57, p=0.003).	Palatoplasties performed by high-volume surgeons are more likely to result in better postoperative outcomes (i.e., lower rates of secondary palatal management) as compared with palatoplasties performed by low-volume surgeons. Case-mix Level 0

Appendix II: Pathways for the Volume/Outcome Relationship



Appendix III: Detailed Quality Table

TOPIC	Year	Reporting	External Validity	Bias	Confounding	Power	Total QI	Level of case-mix
APPENDICITIS								
Chamberlain JM	2000	9	3	5	4	1	22	2
Chamberlain JM	2001	9	3	5	4	1	22	2
median							22	2
CANCER								
Albright AL	2000	8	0	5	3	1	17	0
Danjoux CE	1996	10	3	5	4	1	23	0
Duffner PK	1982	8	3	5	4	1	21	0
Foreman NK	1996	10	3	5	4	1	23	2
Kramer S	1984	10	3	5	4	1	23	0
Meadows AT	1983	10	3	5	4	1	23	0
Sanders BM	1988	7	3	4	4	1	19	2
Stiller CA	1988	7	1	4	3	1	16	1
Stiller CA	1989	10	3	3	4	1	21	1
Stiller CA	1994	8	3	4	4	1	20	2
Stiller CA	1999	11	3	5	4	1	24	0
median							21	0
CARDIAC								
Chameides L	1975	5	3	2	2	0	12	0
Chang R-K	2000	8	3	5	4	1	21	2
Flyer DC	1972	10	3	4	3	0	20	0
Hannan EL	1998	10	3	5	4	1	23	3
Jenkins KJ	1995	10	3	5	4	1	23	2
Lundstrom NR	2000	10	1	4	3	1	19	2
NERICP	1980	5	0	3	3	1	12	2
Norwood W	1988	8	0	4	2	1	15	0
Scott WA	2001	10	3	4	4	1	22	3

Sollano JA	1999	10	3	5	4	1	23	2
Spiegelhalter DJ	2001	8	0	5	4	1	18	2
Stark J	2000	8	3	3	3	1	18	2
median							19.5	2
LEVEL OF CARE								
Baruffi G	1984	11	3	5	4	1	24	2
Becerra JE	1989	7	3	5	4	1	20	3
Berg CJ	1984	10	3	5	4	1	23	2
Black DP	1984	9	3	4	3	1	20	0
Bode MM	2001	10	3	5	4	1	23	2
Bowes WA	1981	6	3	5	3	0	17	0
Clarke M	1993	9	1	5	4	1	20	1
Cordero L	1988	6	3	5	3	1	18	2
Field D	1991	9	3	5	3	1	21	2
Gortmaker S	1985	9	3	5	4	1	22	2
Lagoe RJ	1986	6	3	5	4	0	18	0
LeFevre M	1992	9	3	5	4	1	22	2
Linder-Pelz S	1990	9	1	5	4	0	19	2
Mayfield JA	1990	10	3	5	4	1	23	2
Paneth N	1987	7	3	5	4	1	20	2
Roder D	1993	9	3	5	4	1	22	3
Rosenblatt RA	1985	8	3	4	4	1	20	2
Rosenblatt RA	1988	11	3	5	4	1	24	2
Rosenblatt RA	1996	7	3	4	2	0	16	2
Sangala V	1990	8	1	5	4	1	19	1
Verloove-Vanhorick SP	1988	10	3	5	4	1	23	3
Viisainen K	1994	9	3	5	4	1	22	2
Yeast JD	1998	10	3	5	4	1	23	2
median							20	2
NEONATAL								
Bakketeig LS	1978	10	3	5	4	1	23	2
Callon HF	1975	4	3	4	3	0	14	0
Campbell DM	1984	4	0	2	2	1	9	0
Deutchman ME	1995	8	1	5	3	1	18	1

Fallis G	1988	9	0	4	2	1	16	2
Hein HA	1978	8	3	4	3	1	19	0
Hein HA	1980	9	3	5	3	0	20	2
Hein HA	1981	8	1	5	1	0	15	2
Hein HA	1986	8	3	5	3	1	20	2
Hein HA	1999	7	3	5	3	0	18	2
Hoekstra R	1981	5	1	4	2	1	13	1+
International Neonatal Network	2000	11	0	5	4	1	21	2
Klein MC	2002	10	3	5	4	1	23	2
LeFevre ML	1992	9	3	5	4	0	21	1
Lumley J	1992	7	3	5	4	1	20	2
McCormick MC	1985	9	3	5	4	1	22	2
Moster D	1999	9	3	5	4	1	22	2
Moster D	2001	10	3	5	4	1	23	2
Peddle LJ	1983	4	3	5	3	1	16	0
Piekkala P	1985	9	3	4	4	0	20	2
Rozier JC	1980	6	3	4	3	0	16	0
Shapiro S	1982	6	0	4	2	0	12	2
Siegel E	1985	10	3	5	4	1	23	2
Siegel E	1986	9	3	5	4	1	22	2
Taylor J	1989	7	3	5	3	1	19	0
Usher R	1977	5	3	4	3	0	15	2
Walker DJ	1985	6	3	5	3	1	18	2
Williams RC	1979	8	3	5	4	1	21	2
median							20	2
NICU								
Field D	1999	11	3	5	4	1	24	2
Goodman DC	2002	0	0	0	0	0	22	2
Horbar JD	1988	9	0	5	3	0	17	0
Horbar JD	1997	8	0	5	4	0	17	2
Intern Neonatal Network	2000	11	0	5	4	1	21	3
Paneth N	1982	10	3	5	4	1	23	2
Phibbs CS	1996	9	2	4	4	1	20	2
UK Neonatal Group	2002	10	1	3	3	2	19	2
median							20.5	2

PICU							
DeNicola LK	1992	5	3	3	3	0	14	0
Lopez-Herce J	2000	5	0	3	3	0	11	0
Pearson G	1997	11	3	4	4	1	23	3
Pearson G	2001	9	3	5	3	1	21	0
Pollack MM	1991	11	3	5	4	1	24	2
Pollack MM	1994	10	3	5	3	0	21	2
Tilford JM	2000	10	1	5	4	1	21	2
median							21	2
TRANSPLANTS							
Cox K	2001	8	3	5	4	1	21	1
Hosenpud JD	1994	10	3	5	4	1	23	0
Schurman SJ	1999	10	3	5	4	1	23	2
median		10	3	5	4	1	23	1
TRAUMA							
Cooper A	1993	8	3	4	3	1	19	2
Cosentino CM	1991	10	3	4	4	1	22	0
Konvolinka CW	1995	7	3	5	3	0	18	2
Nakayama DK	1992	10	3	4	4	1	22	2
Potoka DA	2000	10	3	5	4	1	23	2
Tepas JJ	1998	10	3	5	4	0	22	2
Van Rooyen MJ	1995	9	3	4	4	0	20	0
median							22	2
MISCELLANEOUS							
Albright AL	1999	7	3	5	3	0	18	0
Baumer JH	1997	10	3	5	4	1	23	2
Bratton SL	2001	10	3	5	3	1	22	2
Snow BW	1996	9	3	5	3	1	21	0
Witt PD	1998	7	0	5	3	1	16	0
median							21	0

Appendix IV: Literature Search Strategies

The search strategies were based on the following questions:

1) What benefit do higher volume services in pediatrics offer in terms of greater expertise and better outcomes
2) What benefit do higher volume services offer in terms of greater efficiency

The following databases were searched: MEDLINE (1966-October 2002), EMBASE (1988 to October 2002), Science Citation Index (1975 to October 2002), Health Star (1975 to September 2002) and CINAHL (1982 to October 2002). PubMED was also searched in October 2002 to capture any recently published studies.

Due to the lack of standard indexing for this topic, very broad search strategies were used. The extremely high yields of these searches were reduced from approximately 17,000 to 2,700 records by limiting to children. Publication types were not limited to experimental study designs and included observational studies. Searches were not restricted to English.

Approximately 80 of the total included studies for this review were found through the electronic searches and all identified in MEDLINE and EMBASE. No unique studies were found in any of the other four databases. The remaining studies were identified through reference lists and contacting authors of relevant studies and experts in the field.

The same search strategy was used for HealthStar as MEDLINE and the PUBMED strategy consisted of the first five lines of the MEDLINE search. These are therefore not included as separate strategies in the appendix.

Database: MEDLINE <1966 to October 2002>
Search Strategy:

```
01  ((surg$ or hospital$ or centre$ or center$ or facilit$ or institution$ or
    provider$ or procedur$) adj (volume or volumes or caseload$ or
    workload$)).ti,ab.
02  ((physician$ or doctor$ or staff or specialist$ or surg$ or nurs$ or
    cardiologist$ or oncologist$ or practitioner$) adj (caseload$ or
    workload$)).ti,ab.
03  (volume adj2 (outcome or coutomes or mortality or cost$ or length of
    stay)).ti,ab,sh.
04  (patient$ adj (volume or volumes or caseload$)).ti,ab.
05  (procedur$ adj frequen$).ti,ab.
06  or/1-5
07  Specialities, Medical/
08  exp Regional Health Planning/
09  Delivery of Health Care, Integrated/
10  or/7-9
11  "outcome and process assessment"/
12  exp "outcome assessment (health care)"/
13  exp utilization review/
14  exp "task performance and analysis"/
15  exp "costs and cost analysis"/sn
16  or/11-15
17  exp mortality/
18  exp hospitalization/ec, og, sn, ut
19  exp hospital mortality/
20  exp Postoperative complications/
21  exp Survival analysis/
22  exp Patient acceptance of health care/
23  or/17-22
24  exp workload/
25  (admission adj2 rate#).mp.
26  exp professional practice/ec, ma, sn, ut
27  exp case management/ec, og, sn, ut
28  exp health facilities/ut
29  exp physician's practice patterns/
30  or/24-29
31  exp quality of health care/ec, og, sn, ut
32  16 or 31
33  exp *centralized hospital services/ec, ma, og, sn
34  10 or 33
35  exp diagnosis/ut
36  exp surgical procedures, operative/ut
37  exp therapeutics/ut
38  exp patient care/ut
39  or/35-38
40  30 and 39
41  exp clinical competence/
42  or/1-2,4-5,30
43  32 and 42
44  32 and 34
45  23 and 34
46  42 and 41
47  or/3,43-45,46
48  pediatric$.mp.
49  paediatric$.mp.
50  (child$ or newborn$ or adolescen$ or infan$).mp.
51  (preschool$ or nursery school$ or elementary school$ or schoolchild$).mp.
52  (youth$ or teen#).mp.
53  (baby or babies or neonat$).mp.
54  (daycare$ or day care$).mp.
55  or/48-54
```

```
56   exp Child/ or exp Child, preschool/
57   exp Infant/ or exp Infant, Newborn/
58   exp Adolescence/
59   or/56-58
60   47 and 59
```

Database: EMBASE <1988 to October 2002>
Search Strategy:

```
01   exp job performance/
02   exp productivity/
03   or/1-2
04   exp mortality/
05   exp treatment outcome/
06   health care quality/
07   exp patient attitude/
08   or/4-7
09   exp health care facility/
10   exp health service/
11   or/9-10
12   exp centralization/ or exp regionalization/
13   exp *centralization/ or exp *regionalization/
14   (3 or 8 or 11) and 12
15   13 or 14
16   exp workload/ or exp clinical practice/ or exp hospital admission/
17   exp *workload/ or exp *clinical practice/ or exp *hospital admission/
18   3 and 16 (
19   exp *job performance/ or exp *productivity/ or exp *mortality/ or exp
     *treatment outcome/ or *health care quality/ or exp *patient attitude/
20   17 and 19
21   or/15,18,20
22   ((surg$ or hospital$ or centre$ or centre$ or institution$ or facility$
     provider$ or procedur$) adj (volume or volumes or caseload$ or
     workload$)).ti,ab.
23   ((physician$ or doctor$ or staff or specialist$ or surg$ or nurs$ or
     cardiologist$ or oncologist$ or practitioner$) adj (caseload$ or
     workload$)).ti,ab.
24   (volume adj2 (outcome or outcomes or mortality or cost$ or length of
     stay)).ti,ab,sh. (
25   (patient$ adj (volume or volumes or caseload$)).ti,ab.
26   (procedur$ adj frequen$).ti,ab. (
27   or/22-26
28   21 or 27
29   pediatric$.mp.
30   paediatric$.mp.
31   (child$ or newborn$ or adolescen$ or infan$).mp.
32   (preschool$ or nursery school$ or elementary school$ or schoolchild$).mp.
33   (youth$ or teen#).mp.
34   (baby or babies or neonat$).mp. (
35   (daycare$ or day care$).mp.
36   or/29-35 (
37   exp Child/ (
38   exp Infant/ or exp Infant, Newborn/
39   exp Adolescence/
40   exp Child, preschool/
41   or/37-40
42   36 or 41
43   28 and 42
```

Database: CINAHL <1982 to November 2002>
Search Strategy:

---

01 hospital volume.mp.
02 (volume adj2 (outcome# or mortality or morbidity or cost# or admission or efficien$ or length of stay)).mp.
03 or/1-2
04 exp *productivity/
05 exp hospitals/
06 provider experience.mp.
07 physician experience.mp.
08 surgical volume.mp.
09 patient volume.mp.
10 physician workload#.mp.
11 physician$ caseload$.mp.
12 patient caseload$.mp.
13 centre volume.mp.
14 exp workload/
15 or/4-5
16 or/3,6-14
17 15 and 16
18 volume.mp.
19 15 and 18
20 or/17,19

WEB OF SCIENCE (1975 to October 2002)

#1 TI=((physician* OR doctor* OR staff OR specialist* OR surgical OR surgeon* OR nurs*) SAME (caseload* OR workload*)))

#2 TI=(volume SAME (outcome* OR mortality OR morbidity OR cost* OR admission OR efficien* OR length of stay))

#3 TI=((patient OR surgical OR hospital* OR centre* OR centre* OR institution* OR provider*) SAME (volume OR volumes))

#4 #1 or #2 or #3

Limited to Doc Types: Article, Abstract of Published Item, Book Review, Discussion, Meeting Abstract, Meeting-Abstract, Review

Languages: All

Appendix V: Inclusion Form

Date: dd _____ mm _____ 200_____ **Reviewer:** _____ **Reference Number:** _____

1. TOPIC, included if:

[ ] primary research comparing hospitals/centres with different volume levels or before and after concentration of services within a hospital/centre, in terms of patient outcomes (e.g., morbidity, mortality), professional performance (e.g., types of procedures, prescribing patterns), costs, or other measures of efficiency (e.g., LOS)

2. DESIGN, included if:

[ ] study must include a comparison group (RCT, controlled trial, before/after, prospective cohort, retrospective cohort, case control, cross-sectional)

3. PARTICIPANTS, included if:

[ ] one of the following: facilities delivering child health; health workers delivering child care; users of pediatric health services; all providers of child health including governmental (either free health services or systems of cost recovery), non-governmental organizations or the private sector

4. INTERVENTION, included if:

[ ] any pediatric health service

5. OUTCOME, included if:

[ ] studies report at least one objective, quantified outcome (e.g., morbidity, mortality, length of stay, admission rates, service costs, adverse outcomes, user acceptability such as travel costs, expressed satisfaction)

[ ] INCLUDE (meets all inclusion criteria above)

[ ] EXCLUDE – state reason: _____

[ ] CAN'T TELL (need more information)

If disagreement between reviewers, final outcome:

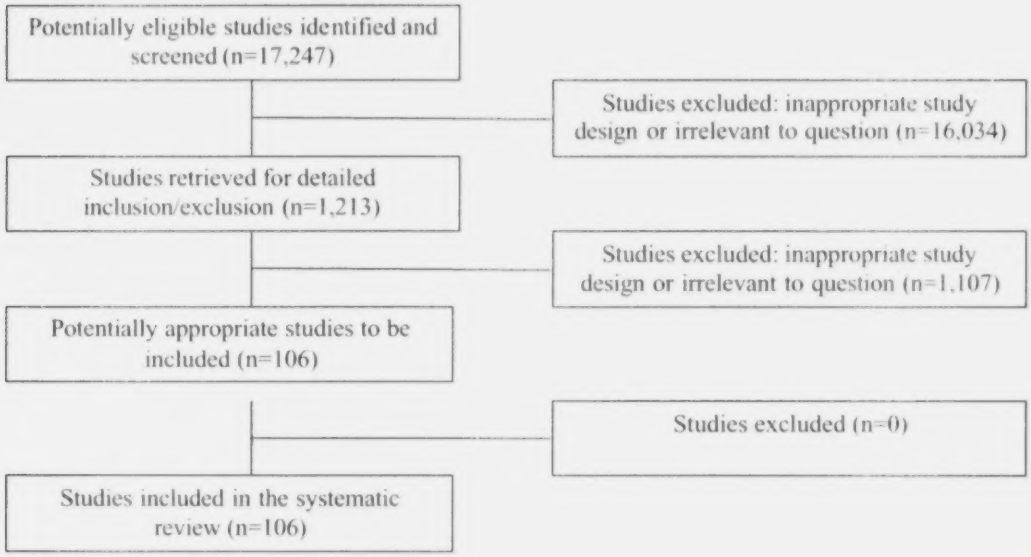
[ ] INCLUDED [ ] EXCLUDED

Diagnosis examined: _____

Procedure examined: _____
Other : _____

Check box if study provides useful background information

Appendix VI: Figure of Inclusion/Exclusion Flow Chart


Potentially eligible studies identified and screened (n=17,247)
Studies excluded: inappropriate study design or irrelevant to question (n=16,034)
Studies retrieved for detailed inclusion/exclusion (n=1,213)
Studies excluded: inappropriate study design or irrelevant to question (n=1,107)
Potentially appropriate studies to be included (n=106)
Studies excluded (n=0)
Studies included in the systematic review (n=106)

Appendix VII: Down's and Black Quality Assessment

Quality Assessment Form for Non-Randomized Controlled Trials

Reporting

1. **Is the hypothesis/aim/objective of the study clearly described?** This question refers to a clear statement of the objective, i.e. to measure the effectiveness of x in population y with respect to z, even if x, y and z are not clearly described (see questions 2, 3 and 4)

Yes	1
No	0

2. **Are the main outcomes to be measured clearly described in the Introduction or Methods section?** If the main outcomes are first mentioned in the Results section, the question should be answered no. In case-control studies the case definition should be considered the outcome.

Yes	1
No	0

3. **Are the characteristics of the patients included in the study clearly described in the Introduction or Methods section?** In cohort studies and trials, inclusion and or exclusion criteria should be given. In case-control studies, a case definition and the source for controls should be given.

Yes	1
No	0

4. **Are the interventions of interest clearly described in the Introduction or Methods section?** Treatments and placebo (where relevant) that are to be compared should be clearly described.

Yes	1
No	0

5. **Are the distributions of principal confounders in each group of subjects to be compared clearly described?** A list of principal confounders is provided.

Yes	2	
Partially	1	
No	0	

6. **Are the main findings of the study clearly described?** Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. This question does not cover statistical tests, which are considered below.

Yes	1
No	0

7. **Does the study provide estimates of the random variability in the data for the main outcomes?** In non-normally distributed data the inter-quartile range of results should be reported. In normally distributed data the standard error, standard deviation or confidence intervals should be reported. If the distribution of the data is not described, it must be assumed that the estimates used were appropriate and the question should be answered yes.

Yes	1
No	0

8. **Have all important adverse events that may be a consequence of the intervention been reported?** This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).

Yes	1
No	0

9. **Have the characteristics of patients lost to follow-up been described?** This should be answered yes where there were no losses to follow-up or where losses to follow-up were so small that findings would be unaffected by their inclusion. This should be answered no where a study does no report the number of patients lost to follow-up.

Yes	1
No	0

10. **Have 95% CIs and/or actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?** (both CI and p value, either CI or p value, neither)

Yes	1
No	0

External validity

11. **Were the subjects asked to participate in the study representative of the entire population from which they were recruited?** The study must identify the source population for patients and describe how the patients were selected. Patients would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample. Random sampling is only feasible where a list of all members of the relevant population exists. Where a study does not report the proportion of the source population from which the patients are derived, the question should be answered as unable to determine.

Yes	1	
No	0	
Unable to determine	0	

12. **Were those subjects who were prepared to participate representative of the entire population from which they were recruited?** The proportion of those asked who agreed should be stated. Validation that the sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.

Yes	1	

No	0	
Unable to determine	0	

13. **Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?** For the study to be answered yes the study should demonstrate that the intervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.

Yes	1	
No	0	
Unable to determine	0	

Internal validity – bias

14. **Was an attempt made to blind study subjects to the intervention they have received?** For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.

Yes	1	
No	0	
Unable to determine	0	

15. **Was an attempt made to blind those measuring the main outcomes of the intervention?**

Yes	1	
No	0	
Unable to determine	0	

16. **If any of the results of the study were based on "data dredging", was this made clear?** Any analyses that had not been planned at the outset of the study should be clearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.

Yes	1	
No	0	
Unable to determine	0	

17. **In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?** Where follow-up was the same for all study patients that answer should be yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow-up are ignored should be answered no.

Yes	1	
No	0	
Unable to determine	0	

18. **Were the statistical tests used to assess the main outcomes appropriate?** The statistical techniques used must be appropriate to the data. For example non-parametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.

Yes	1	
No	0	
Unable to determine	0	

19. **Was compliance with the interventions reliable?** Where there was non compliance with the allocated treatment or where there was contamination of one group, the question should be answered no. For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.

Yes	1	
No	0	

Unable to determine	0	

20. Were the main outcome measures used accurate (valid and reliable)? For studies where the outcome measured are clearly described, the question should be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.

Yes	1	
No	0	
Unable to determine	0	

Internal validity – confounding (selection bias)

21. Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population? For example, patients for all comparison groups should be selected from the same hospital. The question should be answered unable to determine for cohort and case-control studies where there is no information concerning the source of patients included in the study.

Yes	1	
No	0	
Unable to determine	0	

22. Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time? For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.

Yes	1	
No	0	
Unable to determine	0	

23. **Were the subjects randomized to intervention groups?** Studies which state that subjects were randomized should be answered yes except where method of randomization would not ensure random allocation. For example alternate allocation would score no because it is predictable.

Yes	1	
No	0	
Unable to determine	0	

24. **Was the randomized intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?** All non-randomized studies should be answered no. If assignment was concealed from patients but not from staff, it should be answered no.

Yes	1	
No	0	
Unable to determine	0	

25. **Was the re adequate adjustment for confounding in the analyses from which the main findings were drawn?** This question should be answered no for trials if: the main conclusions of the study were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders different between the treatment groups but was not taken into account in the analyses. In non-randomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.

Yes	1	
No	0	
Unable to determine	0	

26. **Were losses to patients to follow-up take into account?** (yes, no, unable to determine) If the numbers of patients lost to follow-up are not reported, the question

should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.

Yes	1	
No	0	
Unable to determine	0	

Power

27. Was a power calculation reported for the primary outcome?

Yes	1	
No	0	
Unable to determine	0	

28. Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance in less than 5%?

Yes	1	
No	0	
Unable to determine	0	

Appendix VIII: Data Extraction Form

Study Identification	
Author	
Year	
Country	

Objective

Study Design: check the appropriate design	
RCT	
Controlled trial	
Before/after	
Prospective cohort	
Retrospective cohort	
Case control	
Cross-sectional	
Before-after Other	
Explanation of regionalization process:	

What is Volume?

Authors' definition of volume:			
Authors' rationale for volume thresholds:			
Circle type of volume			
Physician	Hospital	Both stratified	Both combined
physicians with different volume	different volumes	Explain:	Explain:
specialist vs. generalist	levels of care		
other	became a regional centre		
	other (ex: beds)		

Is there a minimum volume the HMO's required to cover the procedure? Yes [ ] No [ ]
If yes, is it physician or hospital volume? Hospital [ ] Physician [ ]

Volume Characteristics						
	High volume		Continuum			Low volume
Volume cut-offs						
Hospitals						
Physicians						
Beds						
Patients						
Location						
Years studied						
Number of eligible hospitals						
Number of eligible physicians						

Method for Including Hospitals

Voluntary	
Random sample	
Geographical location (specify)	
Database (specify)	
Other (specify)	

Patient Characteristics						
Diagnosis						
Procedure (elective or emergency)						
Technical or medical management						
	High Volume		Continuum			Low Volume
Number in each cohort						
Age						
Risk score						

Source of Patient Information	
Medical records	
Clinical database/registry	
Administrative database	
Insurance claims forms	
Other (specify)	

Risk Adjustment		
Control for case mix in analysis?	Yes	No
Demographic and administrative variables (specify)		
Co-morbidities (specify)		
Stage/severity of illness (specify)		
Better process of care (specify)		
Other (specify)		

Statistical method used to adjust for case-mix	

Authors' Conclusions		
Endorse regionalization	Oppose regionalization	Unsure
Did authors' conclude a threshold volume?	Yes	No
Authors' overall conclusions:		

Additional Comments

Programme de recherche sur les politiques
en matière de santé
Résumé du rapport final

Titre :	**L'organisation des services de santé pour les enfants et les jeunes : synthèse des données pour guider l'intégration et le regroupement des services de santé pédiatriques**
Nom du chercheur :	**Dr Terry Klassen**
Date de la fin du projet :	**Mai 2003**
Catégorie de recherche :	**Synthèse**
Institution :	**Université d'Alberta**
Numéro du projet :	**6795-15-2001/4450003**

Résumé

Il y a plus de trente ans, la relation entre le volume et les résultats a été présentée pour expliquer les différences entre les résultats pour la santé. La relation entre le volume et les résultats fait référence au concept selon lequel plus une procédure est utilisée ou plus un type de patient est traité, plus le médecin et l'hôpital acquerront des compétences et des connaissances spécialisées dans le traitement. Deux théories ont été élaborées dans le but d'expliquer la relation entre le volume et les résultats. La théorie selon laquelle « à force de forger, on devient forgeron » énonce que la répétition d'une procédure et du traitement d'un certain type de patient améliore les résultats. La théorie du « renvoi sélectif » énonce que les patients ayant fait l'objet de renvois vers d'autres médecins ont de meilleurs résultats, et que plus les renvois de patients augmentent, plus le volume des patients traités par le médecin augmente. Afin de déterminer adéquatement de quelle façon le volume influence les résultats, le risque que les patients obtiennent de mauvais résultats doit être normalisé ou ajusté. Nous avons examiné systématiquement 106 études portant sur la façon dont le volume ou la régionalisation des services de santé avait une incidence sur les résultats pour les patients ou l'utilisation des services de santé.

Un bibliothécaire médical a effectué de recherches électroniques dans les cinq bases de données bibliographiques suivantes : Medline, EMBASE, Cinhal, HealthStar, et Web of Science. Deux examinateurs ont vérifié les résultats de recherche et le texte complet des études potentiellement pertinentes a été obtenu. Deux examinateurs ont appliqué les critères d'inclusion à 1 213 études. Les études étaient incluses si elles portaient sur une

relation entre le volume et les résultats ou si elles présentaient une évaluation de l'incidence de la régionalisation des services chez les populations pédiatriques. L'étude doit avoir porté sur un groupe de référence et mesuré les résultats objectifs tels que la mortalité, la durée de séjour ou les taux d'admission. Ensuite, les participants devaient être des établissements ou des travailleurs de la santé fournissant des services de santé pédiatriques, ou tout autre fournisseur de soins aux enfants (notamment les organisations gouvernementales, non gouvernementales ou privées). Quatre-vingt-sept études ont été jugées pertinentes. Dix autres études ont été issues de travaux de congrès ou de listes d'études incluses. Des lettres à l'intention d'auteurs au sujet d'études incluses ont permis de cerner neuf autres études. La qualité a été évaluée de façon indépendante par deux examinateurs. Deux techniques ont été utilisées pour évaluer la rigueur méthodologique des études : la liste de vérification Downs et Black's pour les études randomisées et les études non randomisées sur les interventions en soins de santé ainsi que le degré d'ajustement du risque pour divers groupes clients. Les données ont été extraites par un premier examinateur; leur précision et leur intégralité ont été évaluée par un deuxième examinateur. En raison de la grande hétérogénéité, il n'a pas été possible d'entreprendre une analyse quantitative.

Les secteurs cliniques suivants ont été étudiés : appendicite (n=2), cancer (11), soins cardiaques (12), niveau de soins (23), soins néonatals (28), unité néonatale des soins intensifs (8), unité de soins pédiatriques intensifs (7), greffes (3), traumas (7), et divers (5). L'indice de qualité moyen était de 21 (écart interquartile de 18,22), et la majorité des études ont été entreprises aux États-Unis. La qualité des études s'est améliorée au fil du temps.

Les éléments probants associés à la relation entre le volume et les résultats pour la gestion de l'appendicite ont été inconcluants. Le taux de survie des enfants atteints d'un cancer a tendance à augmenter lorsque les enfants sont traités dans des centres d'oncologie à haut volume; toutefois, cela dépend du type de cancer étudié. Par ailleurs, les enfants sont susceptibles d'obtenir de meilleurs résultats lorsqu'ils sont traités dans des centres d'oncologie pédiatriques plutôt que dans d'autres types d'institutions. Le volume est également associé positivement avec les résultats chez les enfants ayant besoin d'une transplantation. Une relation entre le volume et les résultats existait chez les enfants devant subir une série de chirurgies cardiaques; la régionalisation des soins cardiaques améliorait également les résultats pour la santé. Un niveau accru de soins a donné lieu a un accroissement du taux de survie en ce qui concerne l'insuffisance de poids à la naissance et les nouveau-nés à risque. En ce qui concerne les unités néonatales des soins intensifs et les unités de soins pédiatriques intensifs, le volume n'était pas lié à l'obtention de meilleurs résultats. Par contre, les patients traités dans des institutions ayant davantage de ressources obtenaient de meilleurs résultats que leurs homologues traités dans des hôpitaux de base. On a découvert que le volume n'améliorait pas les résultats dans les milieux de traumatologie, mais on ne sait pas si la disponibilité accrue des ressources ou le traitement dans des centres de traumatologie pédiatriques a une incidence sur les résultats pour les patients.

Une relation entre le volume et les résultats semble exister dans la plupart des milieux examinés, mais elle semble dépendre de la procédure utilisée ou de la spécialité pédiatrique. Les enfants qui risquent grandement de mourir semblent mieux s'en sortir lorsqu'ils sont traités dans des hôpitaux à haut volume ou de haut niveau. La population pédiatrique gravement atteinte (p. ex. admission à l'unité de soins intensifs) ou encore les enfants gravement malades ou nés avant terme obtiennent de meilleurs résultats lorsqu'ils sont traités dans des hôpitaux à forte intensité de ressources. Il y a également une relation entre le volume et les résultats pour la plupart des procédures de chirurgie pédiatrique examinées.

En fonction des éléments probants examinés, le résultat des procédures chirurgicales est généralement plus élevé dans les hôpitaux à haut volume. Toutefois, cela ne s'applique pas à toutes les procédures chirurgicales, et c'est davantage évident en ce qui concerne les opérations rares et complexes. Puisque la disponibilité des ressources semble être un prédicteur des résultats pour la santé plus important que le volume, les enfants gravement atteints s'en sortent mieux lorsqu'ils sont traités dans des institutions qui ont un niveau accru de ressources. Enfin, le résultat des grossesses à haut risque et des enfants nés avant terme et plus élevé lorsque les soins sont prodigués dans des hôpitaux de niveau III. En ce qui concerne la gestion des traumas et des accouchements, un système de transport au point est nécessaire pour transporter les patients vers l'hôpital qui répond à leurs besoins en matière de ressources.

Les opinions exprimées ici ne représentent pas nécessairement le point de vue de Santé Canada

---

En plus de ce résumé, vous pouvez obtenir le rapport des deux manières décrites ci-dessous.

- La version papier du rapport final complet est disponible dans la langue de soumission à la bibliothèque de Santé Canada, par prêt entre bibliothèques.
- La version électronique du rapport final est disponible sur demande dans la langue de soumission. Pour l'obtenir, écrivez à Santé Canada à l'adresse suivante : rmddinfo@hc-sc.gc.ca

Cette recherche a été effectuée grâce à la contribution financière du Programme de recherche sur les politiques en matière de santé de Santé Canada. Pour obtenir la permission de reproduire la totalité ou une partie du rapport de recherche, veuillez communiquer directement avec le chercheur principal à l'adresse suivante : terry.klassen@ualberta.ca

Le Programme de recherche sur les politiques en matière de santé (PRPS) finance des recherches qui fournissent une base de données probantes pour les décisions stratégiques de Santé Canada. Le PRPS est un programme stratégique ciblé ayant

une vaste orientation socioéconomique ainsi que des connexions avec les engagements aux échelons nationaux et internationaux. La recherche peut être primaire, secondaire ou de synthèse, une contribution unique à une initiative de recherche de développement ou bien un atelier, un séminaire ou une conférence

Les détails relatifs au PRPS, ses processus, ses procédures, et son financement, sont accessibles à l'adresse suivante :

http://www.hc-sc.gc.ca/iacb-dgiac/arad-draa/francais/dgdr/prpsindex.html